2011

Helper phage capsid size redirection by staphylococcal pathogenicity island SaPI1 involves internal scaffolding proteins.

Priyadarshan Damle
Virginia Commonwealth University

HELPER PHAGE CAPSID SIZE REDIRECTION BY STAPHYLOCOCCAL PATHOGENICITY ISLAND SaPI1 INVOLVES INTERNAL SCAFFOLDING PROTEINS

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy at Virginia Commonwealth University

by

DAMLE PRIYADARSHAN KAMALKANT

M.Sc., South Gujarat University, 1990

Director: Gail E. Christie, Ph.D.

Professor, Department of Microbiology and Immunology

Virginia Commonwealth University

Richmond, Virginia

Virginia Commonwealth University

Richmond, Virginia

July 2011

Acknowledgement

I would like to thank my mentor Dr. Gail E. Christie for her constant support, guidance, utmost patience and accessibility, whenever I sought her advice. I would also like to thank my committee members Dr. Peterson, Dr. McVoy, Dr. Wright, and Dr. Jefferson for their advice on my research project and progress. I am grateful to Dr. McVoy for his advice, even out of the sphere of this work, and for allowing me to use his laboratory equipment. I am grateful to Dr. Wright and Dr. Peterson for generating an interest in structural biology and enzymology in me and always being there to solve my problems. I am grateful to Dr. Jefferson for agreeing to take Dr. Nicola's place on my committee at a very short notice when he moved out of VCU. I am also thankful to her for her assistance in addressing research problems and allowing me to use her laboratory. I would also like to thank Dr. Nicola for serving on my committee during his tenure at VCU. I thank Dr. Terje Dokland for his advice and guidance as well as the members of his lab; Dr. Anton Poliakov for mass spectrometry and Michael Spilman and Altaira Dearborn for cryo-electron microscopy and 3-D reconstruction. Finally, I would also like to thank the past and present members of Dr. Christie's lab, for their moral support and help.

Dedication

My deepest gratitude goes to my family for their constant love and support in my life. Studying at Virginia Commonwealth University would not have been possible without their inspiration and sacrifices. I am indebted to my parents, Kamalkant Damle and Sujata Damle, for instilling in me the love for education. I am also grateful to my wife, Anita for her unwavering support. She has been a source of inspiration in continuing my education at VCU. I also dedicate my work to my son, Abhishek for being so understanding and never complaining about not being able to spend much of my time with him. I dedicate my work to my family and friends without whom this journey would not have been possible.

Table of Contents

List of Tables

List of figures

List of abbreviations

A	adenine
a.a.	amino acid
att	attachment site
BHI	brain heart infusion
bp	base pair
BSA	bovine serum albumin
BGG	bovine gamma globulin
C	cytosine
CDS	coding sequence or coding region
cfu	colony forming units
cm	centimeter
CsCl	cesium chloride
CSPD	disodium-3-(4-methoxyspiro[1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.13,7]decan)-4-yl)phenyl phosphate
CY-GL	casamino acids yeast extract glycerophosphate
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DIG-11-dUTP	digoxigenin-11-2'-deoxy-uridine-5'-triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dTTP	deoxythymidine triphosphate
dUTP	deoxyuridine triphosphate
EDTA	ethylenediamine tetraacetic acid
E. coli	*Escherichia coli*
ERP	excision replication and packaging
g	gram
G	guanine
gDNA	genomic DNA
gp	gene product
GST	glutathione-S-transferase
HFT	high frequency transduction
IPTG	isopropyl-D-1-thiogalactopyranoside
Jcm	joules per centimeter
kbp	kilobase pairs
L	liter
LB	Luria-Bertani
M	molar
ml	milliliter

mmol	millimole
mM	millimolar
M.O.I.	multiplicity of infection
mol	mole
mRNA	messenger RNA
msec	millisecond
MW	molecular weight
MWCO	molecular weight cut-off
n	nano
NCTC	National Collection of Type Cultures
nm	nanometer
˚C	degree Celsius
OD	optical density
orf	open reading frame
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFU	plaque forming units
PLP	procapsid-like particle
PMF	peptide mass fingerprinting
RBS	ribosome binding site
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal RNA
S. aureus	*Staphylococcus aureus*
SaPI	*Staphylococcus aureus* pathogenicity island
SaPI1	*Staphylococcus aureus* pathogenicity island one
SaPIbov	*Staphylococcus aureus* pathogenicity island bovine isolate
SDS	sodium dodecyl sulfate
sir	SaPI1 Interference Resistant
sbr	SaPIbov Interference Resistant
SSC	saline sodium citrate buffer
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TSB	tryptic soy broth
TSS	transcriptional start site
tRNA	transfer RNA
TU	transducing units
U	uracil
UV	ultraviolet
wt/vol	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-!-D-galactoside
µg	microgram
µm	micrometer
µM	micromolar

Abstract

HELPER PHAGE CAPSID SIZE REDIRECTION BY STAPHYLOCOCCAL PATHOGENICITY ISLAND SaPI1 INVOLVES INTERNAL SCAFFOLDING PROTEINS

by

DAMLE PRIYADARSHAN KAMALKANT, Ph.D

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2011

Major Director: Gail E. Christie, Ph.D.

Professor, Department of Microbiology and Immunology

Virginia Commonwealth University

Richmond, Virginia

Staphylococcus aureus is one of the leading causes of nosocomial and community acquired infections. Many of the virulence factors are encoded on staphylococcal mobile genetic elements. Members of the SaPI family of *S. aureus* mobile elements encode superantigens and are mobilized at high frequency by specific helper bacteriophages. One remarkable feature of helper phage exploitation by SaPIs is remodeling of the normal *T=7* bacteriophage capsid to produce smaller *T=4* phage-like particles. These

particles, composed entirely of helper phage proteins, can accommodate the smaller SaPI genome while excluding that of a complete helper phage. This study was designed to understand the mechanism of capsid size redirection and high frequency mobilization of SaPIs. A multipronged approach employing cryo-EM analysis, protein profile comparison and genetic analysis was used to study the capsid size redirection. Two proteins encoded by the prototype element SaPI1, gp6 and gp7, have been identified in SaPI1 procapsids but not in mature SaPI1 particles. These proteins are sufficient and required to direct small capsid formation, which involves alteration of an internal scaffold. While many phages use internal scaffolding proteins, the involvement of an internal scaffold in capsid size redirection is novel.

Chapter 1

Introduction

S. aureus:

Taxonomic classification:

Domain: Bacteria
Kingdom: Eubacteria
Phylum: Firmicutes
Class: Bacilli
Order: Bacillales
Family: Staphylococcaceae
Genus: *Staphylococcus*
Species: ***Staphylococcus aureus***

Staphylococcus aureus is a Gram positive facultative anaerobe that grows in golden yellow grape-like clusters. They form grape-like clusters because of their ability to divide in two planes in contrast to streptococci which can divide in a single plane to form chains. The distinct golden yellow color is due to a membrane bound triterpenoid carotenoid pigment Staphyloxanthin. This pigment is an anti-oxidant that plays a role in the fitness of the species by neutralizing reactive oxygen species (ROS) generated by the host immune system (Clauditz *et al.*, 2006). *S. aureus* is catalase positive and can convert hydrogen peroxide to water and oxygen. It is also coagulase positive and oxidase negative. All these characteristics help in readily identifying and differentiating these microbes.

Epidemiology of *Staphylococcus aureus*

Approximately 20-25% of humans are persistent carriers, 55-60% are intermittent carriers and 20% are non-carriers (Vandenberg *et al.*, 1999). Nearly all humans suffer from staphylococcal infections at some point in their life time. Although most of the time *S. aureus* causes minor infections ranging from pimples, boils, conjunctivitis and uncomplicated wound infections, it is a leading cause of community acquired pneumonia and can also cause sepsis, toxic shock syndrome, osteomyelitis, meningitis, and fulminant illnesses such as acute bacterial endocarditis (ABE) in previously healthy individuals (Foster,1996; Lowy,1998; Kim *et al.*, 2003; Lin *et al.*, 2010; Naughton *et al.*, 2011; Yamada *et al.*, 2011).

Patients in a hospital setting are at risk from many species of microbes and historically *S. aureus* has always been considered one of the most important. One Boston study in the year 1941 found that 82% of the 122 patients treated for *S. aureus* bacteremia died because of their infection (Skinner *et al.*, 1941). Though one could argue that during that period even simple surgical procedures were often life threatening, patients in the modern hospital environment are still at risk from *S. aureus* infections due to more adventurous surgeries in an aging population and the prevalence of antibiotic resistant *S. aureus*. During the 1950s, a *S. aureus* strain known as phage type 80/81 emerged. It was resistant to penicillin and caused serious nosocomial and community-associated infections worldwide. This strain was largely eliminated after the introduction of methicillin and its derivatives in 1959 (Varga *et al.*, 1961). However, within two years, emergence of methicillin resistant *S. aureus* (MRSA) was reported and since then it has spread worldwide. *S. aureus* is the leading cause of hospital-

associated infections and most of them are caused by MRSA. The mortality rates associated with invasive MRSA are as high as 20% (Cosgrove *et al.*, 2003). Outside the hospital setting, there has been a rapid recent emergence of the more virulent community-associated MRSA (CA-MRSA) (Deleo *et al.*, 2010) and it is surprising that these antibiotic resistant strains emerged in a niche that does not have a high selective pressure of antibiotics. CA-MRSA was first reported in 1990 in Western Australia. In the US, the first documented cases of CA-MRSA were reported in 1999. However, retrospective studies have shown occurrence of MRSA as early as 1993 (Witte *et al.*, 1997). Most cases of CA-MRSA in the US are caused by strain USA300, which carries SCC*mec*typeIV, PVL, PSM-α and enterotoxins Q and K and USA400 (MW2), which is a highly virulent CA-MRSA strain (Baba *et al.*, 2002). USA 400 contains *mec* and *blaZ* antibiotic resistance genes, the PVL operon (Voyich *et al.*, 2006; Kernodle,2007), and 16 unique superantigen genes that include 11 exotoxin genes and five enterotoxin genes (Gordon *et al.*, 2008).USA400 also contains a novel gene cluster which is called "bacteriocin of *S. aureus"* (*bsa*). The *bsa* cluster encodes a potential bacteriocin, or lantibiotic agent, that has been implicated in helping USA400 compete with other colonizing flora and increase the chance of infection with this strain (Baba *et al.*, 2002; Gordon *et al.*, 2008). Other community associated MRSA strains that are prevalent in US are USA500 and USA1000. USA100 and USA200 are more abundant in health care associated infections. MRSA is not only confined to humans but is also a growing concern in veterinary medicine and animal agriculture. Many animal species, including dogs, cats, rabbits, horses, cattle, pigs, poultry, and exotic species are healthy carriers of MRSA as well as infected with it. The true scope of MRSA in animals is not properly

understood at the moment but there appears to be a correlation between MRSA in humans and in animals.

Glycopeptide resistant *S. aureus*

There has been an emergence of *S. aureus* strains that are resistant to other antibiotics as well. Glycopeptides, including vancomycin and also teicoplanin, have been used to treat MRSA infections. Vancomycin was first introduced in 1958 (Fairbrother *et al.*, 1956; Schneierson *et al.*, 1958; Dutton *et al.*, 1959; Wilson,1959; Louria *et al.*, 1961; Kirby,1963). It was thought that vancomycin resistance was unlikely to occur in a clinical setting because resistance to this antibiotic was difficult to induce in vitro and almost forty years had passed without an occurrence of a vancomycin resistant strain. However, this hypothesis proved to be incorrect because vancomycin-intermediate *S. aureus* (VISA) strains were isolated in Japan in 1996. VISA strains were then isolated in many other countries including the USA. Furthermore, to make matters worse, the first fully vancomycin-resistant *S. aureus* (VRSA) clinical isolates in the US were found in 2002 (Centers for Disease Control and Prevention (CDC), 2002). Four clinical VRSA isolates have been reported thus far, all in the USA (Zhu *et al.*, 2008). In addition, these VRSA isolates have a different and potentially much more efficient mechanism for conferring resistance than VISA strains. VRSA strains are thought to acquire additional resistance by conjugal transfer of plasmids containing the *vanA* operon from vancomycin-resistant *Enterococcus faecalis*. Whereas resistance in VISA strains is thought to occur solely because of synthesis of peptidoglycan in larger quantity, resistance in VRSA strains is caused by alteration of the cell wall terminal

peptide D-alanyl-D-alanine to D-alanyl-D-lactate. This prevents the inhibition of cell-wall synthesis by vancomycin.

The spread of antibiotic resistant *S. aureus* strains is just the tip of an iceberg because of the versatility of this microbe owing to the vast array of its virulence factors. These virulence factors have aided this microorganism to adapt in various environments and evade the host immune mechanisms, and thereby have made it a pestilent organism. Stated below is a brief summary of the important virulence factors that are carried by *S. aureus*.

Staphylococcal arsenal of virulence factors

S. aureus colonizes animals and is common in mammals. The primary niche of this organism in humans is *vestibulum nasi* but it is occasionally also found in the armpit, groin, digestive and vaginal tract (Smith *et al.*, 1982). The ability of *S. aureus* to survive in diverse environments is ascribed to its capacity to modulate behavior in response to changes in its environment. It can sense oxygen reduction through the activity of SrrAB and can switch from a respiratory pathway to a fermentative pathway using this two component system (Throup *et al.*, 2001; Yarwood *et al.*, 2001). *S. aureus* has evolved to sense intracellular concentrations of various metal ions through metalloregulatory proteins like Fur, Zur and MntR (Horsburgh *et al.*, 2001; Lindsay *et al.*, 2001). It also overcomes nitrosative and oxidative stress via the SrrAB and Fur regulons (Richardson *et al.*, 2006).

S. aureus has evolved complex mechanisms for causing infection in different environments. Production of virulence factors is necessary to establish infection. These virulence factors are controlled by two broad categories of regulatory systems, namely

the two-component signal transduction system and global transcription regulation. The best studied two-component regulatory system in *S. aureus* is the *agr* (accessory gene regulatory) locus. It regulates more than 70 genes, including 23 virulence factors (Ziebandt *et al.*, 2004). The *sae* (staphylococcal accessory element) locus encodes yet another two-component system SaeRS, which is involved in regulation of many virulence factors. SaeR positively regulates α, β, and γ-hemolysins (Goerke *et al.*, 2005) and negatively regulates protein A (Giraudo *et al.*, 1997). Inactivation of the *sae* locus has been demonstrated to decrease virulence in animal infection models (Novick, 2003; Goerke *et al.*, 2005). Additionally, several global transcriptional regulators also play a role in regulation of virulence factors (Cheung *et al.*, 2008). SarA (Staphylococcal accessory regulator) upregulates expression of FnBPs (fibronectin binding proteins) and α-toxin and decreases expression of protein A and proteases. Apart from playing a role in the overall general stress response, the alternate sigma factor σ^B, plays a central role in cellular physiology and transcription of virulence associated genes. σ^B regulates at least 30 virulence genes, including *sar*A.

There are three essential steps involved in establishing an infection: (1) cell surface binding, (2) host immune system evasion, and (3) dissemination and tissue invasion. Virulence factors are hence classified according to their role in the infection process. Some of these virulence factors are bound to the cell surface while others are secreted into the extracellular medium. Cell wall associated factors include exopolysaccharides, cell wall components such as teichoic acid and peptidoglycans, and two broad categories of adhesins (Clarke *et al.*, 2006), namely the MSCRAMMS (microbial surface components recognizing adhesive matrix molecules) and SERAMS

(secreted expanded-repertoire adhesive molecules). MSCRAMMS are cell wall associated proteins that are covalently linked to peptidoglycan and aid the bacterium in attachment to extracellular matrixes such as collagen and fibronectin. They include collagen adhesion (Cna), fibronectin binding proteins (FnBPs), clumping factors A and B (ClfA and ClfB) and protein A (Spa). Fibronectin binding proteins, FnbpA and FnbpB, bind to fibronectin and have been shown to be involved in invasion in some cell lines. Once FnbPs have bound to fibronectin, they also bind to fibronectin receptor-integrin, thus facilitating uptake into the host cell. Clumping factors, ClfA and ClfB, bind to fibrinogen and type I cytokeratin-10, which leads to platelet aggregation. ClfA also protects the bacteria from phagocytosis by macrophages. Protein A interferes with the host immune response by causing B cell sensitization (Bekeredjian-Ding *et al.*, 2007) and binding to antibodies such as IgG, IgA and IgE(O'Brien *et al.*, 2002). SERAMs are secreted molecules that include extracellular fibrinogen binding proteins (Efb), extracellular matrix proteins (Emp) and extracellular adhesive proteins (Eap). These proteins not only help in cell adherence (Haggar *et al.*, 2003; Hauck *et al.*, 2006), but also modulate host immune response (Harraghy *et al.*, 2003).

Another important class of virulence factors is the exopolysaccharides. There are two important types of exopolysaccharides in *S. aureus*, namely the PIA (polysaccharide intercellular adhesions) and the capsular polysaccharides. Capsular polysaccharides protect the bacteria from the host immune system by interfering with phagocytosis (Peterson *et al.*, 1978). PIA is a biofilm exopolysaccharide made up of poly-N-acetyl-glucosamine and plays a role in biofilm formation.

S. aureus also secretes a variety of extracellular proteins. Most important of these are: (1) membrane-active toxins, (2) pyrogenic toxin superantigens (PTSags) and (3) exfoliative toxins (ETs) (Bohach, 2006). The cytolytic toxins include α-hemolysin, β-hemolysin, γ-hemolysin, δ-hemolysin, leukocidin and Panton-Valentine-leukocidin. These membrane-active toxins damage the cytoplasmic membrane of the eukaryotic host cells thereby causing osmotic lysis of the cells (Foster, 2005).

Membrane-active staphylococcal toxins

α -hemolysin, encoded by the *hla* gene, is secreted as a monomer of 293 amino acids and upon binding to the eukaryotic membrane oligomerizes to form a homo-heptameric β-barrel shaped transmembrane pore (Song *et al.*, 1996). It is known to affect several cell types including erythrocytes, mononuclear immune cells, epithelial and endothelial cells and platelets (Menestrina *et al.*, 2001; Bohach, 2006). β-toxin is encoded by the *hlb* gene and shows species-dependent activity. It can lyse sheep, cow and goat but not rabbit erythrocytes. β-toxin is a neutral sphingomyelin phosphodiesterase or sphingomyelinase (SMase). This enzyme is a hydrolase that is involved in hydrolyzing sphingolipid to phosphocholine and ceramide and hence the degree of erythrocyte sensitivity depends upon the sphingomyelin content of the cell membrane. β-toxin in humans has been shown to kill proliferating human lymphocytes (Huseby *et al.*, 2007). With the help of in vitro studies, β-toxin has also been shown to covalently cross-link to itself in the presence of DNA to form an insoluble nucleoprotein matrix and to stimulate biofilm formation in a rabbit model (Huseby *et al.*, 2010). The δ-hemolysin is a 26-amino-acid peptide that is capable of causing membrane damage in a variety of mammalian cells. This small peptide is encoded by the *hld* gene which is

embedded in the *agr* effector RNAIII. This small RNA has been shown to play a regulatory role in the accessory gene regulator (*agr*) system that controls the production of cell-associated and secreted virulence factors. This toxin acts synergistically with the sphingomyelinase β-toxin to enable the phagosomal escape of staphylococci in human epithelial as well as in endothelial cells (Giese *et al.*, 2011).

S. aureus also produces two types of bicomponent toxins, namely γ-hemolysin and Panton-Valentine (PV) leukocidin. Each of these toxins is made as two non-associated secreted proteins called the S and F components (the classification is based on slow- and fast-eluting components in an ion-exchange column) (Cooney *et al.*, 1993; Colin *et al.*, 1994). Almost all *S. aureus* strains produce γ-hemolysin, while PV leukocidin is produced by 2 to 3% of the strains. The S and F components of the classical PVL are LukS-PV and LukF-PV (Prevost *et al.*, 1995a; Bohach, 2006). However, in some strains LukS-PV is replaced by LukM-PV. The S and F components of gamma toxin are HlgA and HlgB, respectively. The PVL structural genes were shown to be present in three staphylococcal temperate phages, namely ΦPVL, ΦSLT and ΦPV83 (Kaneko *et al.*, 1998; Zou *et al.*, 2000; Narita *et al.*, 2001). PVL oligomerizes on the eukaryotic membrane to form a hetero-octameric trans-membrane pore (Aman *et al.*, 2010). It is nonhemolytic but can stimulate and lyse neutrophils and macrophages (Prevost *et al.*, 1995b; Bohach, 2006). Gamma toxin is strongly hemolytic but is remarkably less leukotoxic (Bohach, 2006).

Staphylococcal pyrogenic toxin superantigens (PTASags)

The PTSAgs are a group of exotoxins secreted by *S. aureus*. These include toxic shock syndrome toxin-1 (TSST-1) and other staphylococcal enterotoxins (SEs A-E, G-J,

and the staphylococcal enterotoxin-like (enterotoxicity unproven) toxins (SEl) K-R and U, and the recently identified SEl-U2 and SEl-V) (Fraser *et al.*, 2008). Most staphylococcal Sags are encoded by accessory genetic elements such as prophages, transposons, plasmids, and pathogenicity islands. For example, the pathogenicity island vSaβ of staphylococcal strain N315 carries the genes for SEG, SEL, SEM, SEN, and SEO. The φSa3 prophage of strain MW2 carries genes for SEA, SEG2, and SEK2. TSST-1 is encoded by *tstH* (where H refers to human isolate), which is present on the bacterial chromosome within pathogenicity islands SaPI1 and SaPI2 (Blomster-Hautamaa *et al.*, 1986; Lindsay *et al.*, 1998).

Not all Sags are enterotoxic and therefore a new nomenclature was introduced in order to distinguish those that were enterotoxic from those that remained unconfirmed (SEls) (Lina *et al.*, 2004). There is remarkable sequence variation among members of the staphylococcal superantigen family, with the most distant members, SEB and SEK, having only 15% amino acid identity. TSST-1 is unique because it is able to cross mucosal surfaces (Bohach, 2006) and is the only superantigen known to reactivate bacterial cell wall-induced arthritis (Schwab *et al.*, 1993). The staphylococcal PTSAgs cause several acute or chronic human diseases (Kotzin *et al.*, 1993; Leung *et al.*, 1998). For example, TSST-1 has been found in the kidneys of 18% of victims of sudden infant death syndrome (Newbould *et al.*, 1989). TSST-1-producing *S. aureus* has been isolated from over 60% of patients with Kawasaki syndrome (Leung *et al.*, 1993; Abinun *et al.*, 1994; Curtis *et al.*, 1994; Nishiyori *et al.*, 1994). Kawasaki syndrome is the major cause of acquired heart disease among children in the United States. Two human diseases that are definitely caused by staphylococcal PTSAgs are staphylococcal food

poisoning (SFP) and staphylococcal toxic shock syndrome (TSS). Contamination of food with staphylococcal PTSags causes gastroenteritis that is manifested clinically as emesis with or without diarrhea (Dinges *et al.*, 2000). Signs of systemic toxicity, such as fever and hypotension, are rarely observed in cases of SFP. Furthermore, SFP is a self-limited condition that typically resolves within 24 to 48 h of onset. Staphylococcal TSS occurs in either of the two general forms; menstrual or non-menstrual. Menstrual TSS usually occurs in young women whose mucosal layers of the vagina are colonized by TSST-1 positive *S. aureus*. Interestingly most menstrual staphylococcal TSS strains were found to be clonal and contained a pathogenicity island, SaPI2 (Subedi *et al.*, 2007). The non-menstrual TSS results from infections elsewhere in the body and may be caused by TSST-1, SEB or SEC (Bohach *et al.*, 1990; Bohach, 2006). TSST-1 is translated as a precursor protein with 234 amino acids and secreted after cleavage of a 40-amino-acid signal sequence located at its N-terminus. The mature protein is a single polypeptide chain with a molecular weight of 22 kDa and contains a high percentage of hydrophobic amino acids, yet it is highly soluble in water. The toxin is resistant to heat and proteolysis. Staphylococcal Sags are potent T cell mitogens but exhibit different preferences for MHC class II alleles and also produce distinct TCR Vβ profiles (Fraser *et al.*, 2008).

Staphylococcal mobile genetic elements

Mobile genetic elements (MGE) are discrete DNA segments that encode enzymes that mediate the movement of DNA within genomes (intracellular mobility) or between bacterial cells (intercellular mobility) (Frost *et al.*, 2005). Staphylococcal MGEs include plasmids, transposons (Tn), insertion sequences (IS), bacteriophages, phage

related chromosomal islands, and staphylococcal chromosomal cassettes (SCC), and account for almost 20% of the staphylococcal genome. Staphylococcal MGEs often encode virulence factors or antibiotic resistance genes that confer an advantage to *S. aureus* to adapt and disseminate in its host. In particular, Panton-Valentine leukocidin (PVL), chemotaxis inhibitory protein (CHIP), staphylokinase (Sak), exfoliative toxin (Eta), staphylococcal complementary inhibitor (SCIN) and SEA are encoded by various staphylococcal temperate phages. Superantigens including Toxic shock syndrome toxin (TSST1) and various enterotoxins encoded by SaPIs and resistance to β–lactams including methicillin and oxacillin, tetracycline, erythromycin, and vancomycin, have made *S. aureus* a major pathogen of the developed world.

Pathogenicity islands

Pathogenicity islands are discrete segments of DNA that encode one or more virulence factors and most often have a different G+C content than the rest of the chromosome indicating that they have been acquired by horizontal transfer (Hacker *et al.*, 1997; Schmidt *et al.*, 2004a). They usually range from 10 Kb to 200 Kb in size and are often located adjacent to tRNA genes, which presumably act as targets for integration because of their conserved sequences across genera and their abundance in any given species. Pathogenicity islands are flanked by direct repeats with a perfect or nearly perfect repetition of DNA sequence. Most of the pathogenicity islands across genera seem to have lost mobility due to loss of functional mobility genes while the ones that still retain mobility use either transduction or conjugation for horizontal transfer. Pathogenicity islands are ubiquitous in *S. aureus* and almost all *S. aureus* clinical isolates sequenced thus far contain one or more.

Staphylococcal superantigen pathogenicity islands (SaPIs)

One closely related family of pathogenicity islands found in *S. aureus* has been named Superantigen Pathogenicity Islands (SaPIs). The genomic architecture of these pathogenicity islands is strikingly phage-like and modular. One of the most important characteristics of these pathogenicity islands is that they depend on helper phages for their mobilization. The frequency at which they are transduced, known as high frequency transduction (hft), is about 1000 times higher than that observed for generalized transduction by these helper phages. Their dependence on helper phages for mobilization has prompted them to be designated as Phage-related chromosomal islands (PRCIs) (Novick *et al.*, 2010). SaPI1 is a prototypical member of PRCIs and encodes TSST1 and two other enterotoxins, namely EntQ and EntK. SaPI1 can excise and circularize in the presence of Φ13 and phage 80α but is only mobilized at a high frequency by phage 80α. SaPI1 always integrates at a single chromosomal site (*attc*) near the *tyrB* locus and is always found in the same orientation (Lindsay *et al.*, 1998).

Genomic architecture of SaPIs

All SaPIs share a strikingly similar genomic architecture and contain a conserved core of homologous genes. A comparison between two closely related SaPIs, SaPIbov1, which has been extensively studied using genetic experiments, and SaPI1 is shown in Fig1. The first module contains genes that are involved in integration and transcriptional regulation. The SaPI integrase homolog is at the left end of the genome and located adjacent to *int* are two divergent promoters that regulate expression of two putative transcriptional regulatory proteins, Stl and Str. Both of these proteins have predicted helix-turn-helix motifs suggesting that they can bind to the DNA. The function

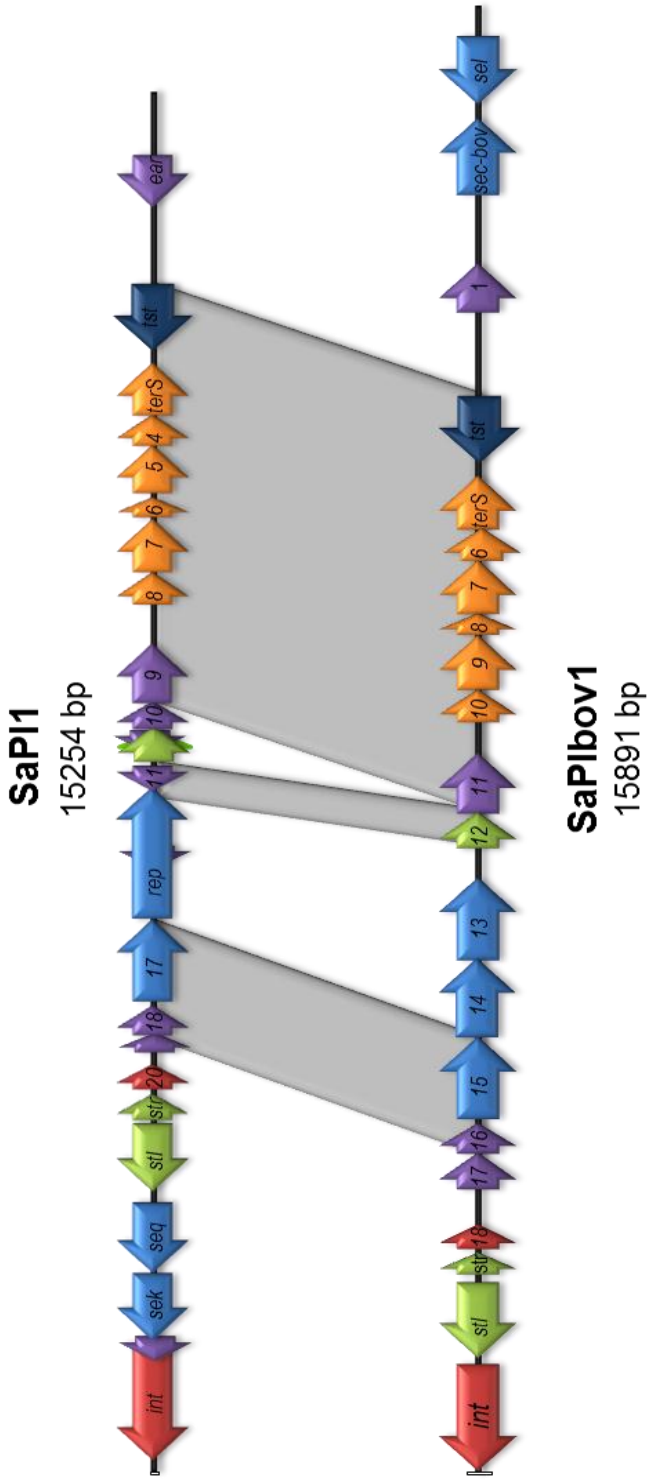
of the Stl regulatory protein bears a close resemblance to cI protein of phage λ. In both SaPIbov1 and SaPI1, Stl is the master repressor that negatively regulates transcription from the *str* promoter (Ubeda *et al.*, 2008). The Stl protein of SaPI1 has similar function as SaPIbov1 Stl though they have no sequence similarity. To the right of *str* lie genes involved in the excision-replication-encapsidation (ERP) cycle.

Excisionase (*xis*) is located immediately downstream of *str* in most of the SaPIs. SaPIs require Int for integrating into the chromosome, while excision is dependent on Xis along with Int. Further to the right of str gene lies the replication module which consists of primase (*pri*), rep helicase (*rep*) and origin of replication (*ori*). The role of the primase is to enhance replication but it is not absolutely essential (Ubeda *et al.*, 2008). The rep helicase is similar to the helicase of lactococcal phage bIL312, and presumably involved in ATP dependent unwinding of double helical DNA to form the ssDNA intermediates required in DNA replication (Korolev *et al.*, 1997). The origin of replication consists of two sets of inversely oriented iteron sequences six to eight bp long and flanking an AT-rich region (Novick *et al.*, 2010). To the right of the origin of replication, and possibly in a separate transcription unit, lie several genes of unknown function. In SaPIbov1, gene12, also called *pif* (phage interference function), is immediately to the right of the origin of replication. The gene product of *pif* was found to interfere with phage 80α growth (Ubeda *et al.*, 2009).

The phage exploitation module lies to the right of the replication module and consists of six genes including a small terminase homolog. These six genes are transcribed from a single promoter, and constitute a single operon known as Operon I.

Figure 1. Comparison of SaPIbov1 and SaPI1 genomes.

Genomes are aligned according to the prophage convention with integrase gene (int) at the left end. Genes are colored according to their sequence and function: int is shown in red; transcription regulators are shown green; replication genes are shown in blue; encapsidation genes are shown in yellow, with the terminase small subunit gene (*terS*) in dark blue; superantigen and other accessory genes are purple. The homologous regions between these two mobile elements are shaded in gray.


SaPI1
15254 bp
ear
tst
terS
4
5
6
7
8
9
10
11
rep
17
18
20
str
stl
seq
sek
int
SaPIbov1
15891 bp
sel
sec-bov
1
tst
terS
6
7
8
9
10
11
12
13
14
15
16
17
18
str
stl
int

Transcription of Operon I is under the control of repressor LexA and is mediated by binding of LexA to two Cheo boxes. These Cheo boxes are found in the 5' region of Operon I and are conserved at the same location in all SaPIs sequenced thus far (Ubeda *et al.*, 2007). Substitution of wt LexA with noncleavable LexA greatly reduced transcription of SaPIbov1 Operon I. Furthermore, deletion of Operon I was found to have no effect on replication of SaPIbov1, but greatly reduced formation of transducing particles (Ubeda *et al.*, 2007). In SaPIbov1, two genes in Operon 1, namely 9 and 8, have been implicated in capsid size redirection. The small terminase homolog is required for SaPI-specific DNA packaging (Ubeda *et al.*, 2007). The roles of the remaining three genes are not clear at present.

SaPI specific accessory genes are located near either or both ends of the SaPI genome. The putative functions of many of these genes have been assigned based on homology or their location in the genome. SaPIs accessory genes identified to date encode various superantigens, toxins, antibiotic resistance and biofilm inducing proteins. SaPI1 carries genes for TSST1 and enterotoxins K and Q while SaPIbov1 carries genes for TSST1 and enterotoxins C and L. Expression of these genes is complex and their regulation is mediated by a range of host and environmental factors. Expression of enterotoxin B has been shown to be inversely related to expression of alternate sigma factor B (Schmidt *et al.*, 2004b) while regulation of *tst* is influenced by environmental conditions such as pH, CO_2 and glucose and is believed to be mediated by catabolite control protein (CcpA) through catabolite responsive (Cre) sites found in the promoter regions of all known *tst* genes (Seidl *et al.*, 2008).

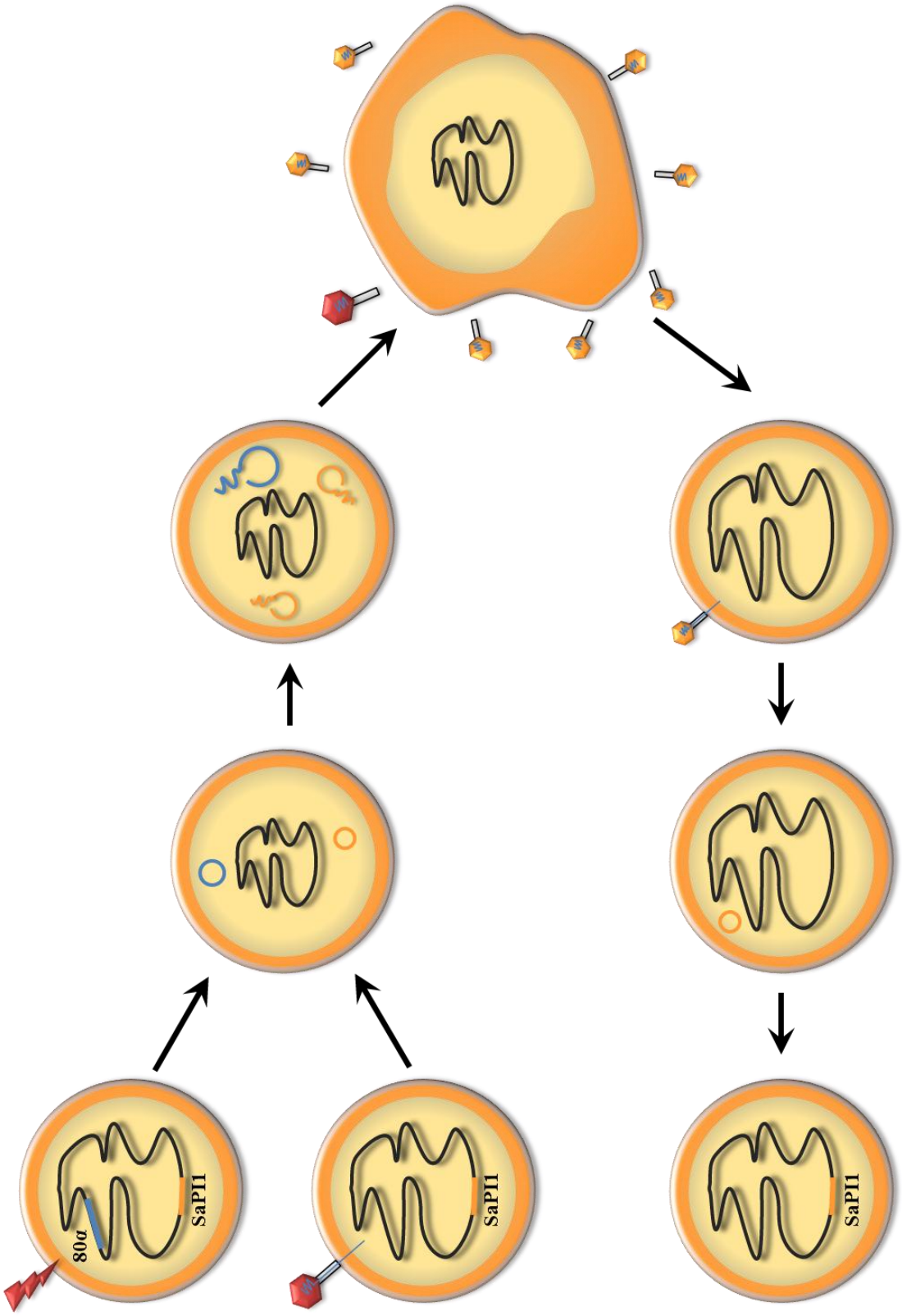
Life cycle of SaPI

SaPIs have a complex life cycle and are dependent on *S. aureus* and helper phage functions for their own propagation (Fig 2). Like prophages, SaPIs are stably integrated into the *S. aureus* chromosome at specific sites. All SaPIs discovered thus far occupy one of the six different *attC* sites in the *S. aureus* genome (Subedi *et al.*, 2007). These *attC* sequences are unique and unrelated to each other. Each SaPI contains a unique insertion site sequence (*attS*) that helps it undergo site specific integration, and no *S. aureus* strains sequenced thus far have been found to contain two SaPIs together at the same location. SaPIs stay stably integrated in the bacterial chromosome because of repression of their divergent promoters by the master regulator Stl. Mutational inactivation of $Stl_{SaPIbov1}$ and Stl_{SaPI1} has been shown to cause spontaneous excision (Tormo-Mas *et al.*, 2010); Bento J.C. unpublished). The site of action of Stl is in the intergenic region between the two divergent promoters for *stl* and *str* where it presumably regulates transcription from both promoters, although only regulation of the rightward *str* promoter has been shown thus far. The role of Str is unclear, because mutational inactivation of Str has no effect on SaPI replication (Ubeda *et al.*, 2008), but *xis* and the replication functions are thought to be expressed from this transcript.

SaPI derepression is not dependent on the SOS response, but requires a helper phage for relief of *stl* repression (Ubeda *et al.*, 2008; Tormo-Mas *et al.*, 2010). Not all SaPIs use the same helper phage proteins for relief of repression. SaPIbov1 containing strains that were defective in helper phage 80α interference were found to contain mutations in the phage encoded *dut* gene, which encodes a bifunctional protein that has

Figure 2. Schematic representation of our model for high frequency mobilization of the SaPIs.

On infection of SaPI positive strain with helper phage 80α, SaPI1 repression is relieved. It excises from the *S. aureus* chromosome and replicates along with the helper phage. SaPI1 DNA is packaged in smaller sized capsids and released along with helper phage after lysis. These SaPI1 particles can infect a SaPI1 negative strain and thereby complete the process of its horizontal transfer.


80α
SaPI1
SaPI1
SaPI1

dUTPase activity and also functions in derepression of SaPIbov1(Tormo-Mas *et al.*, 2010). This protein interacts with $Stl_{SaPIbov1}$ and thereby prevents it from binding to the *stl-str* divergent region. Interestingly, the domain of dUTPase involved in binding to Stl is different from the one involved in dUTPase catalytic activity. Derepression by dUTPase is specific to SaPI1bov1; SaPI1 derepression is caused by a different 80α encoded protein, Sri (Tallent *et al.*, 2007). This protein has also been shown to block bacterial DNA replication (Liu *et al.*, 2004). In a co-purification study, 80α encoded Sri was demonstrated to co-purify with Stl_{SaPI1}, suggesting that like dUTPase this protein directly interacts with its target Stl to block repression (Harwich, 2009)

SaPI derepression, following induction of a resident helper prophage by the SOS response or by superinfection with a helper phage, leads to excision of SaPI1 from the chromosome and replication as a high molecular weight concatemer. During replication, SaPI1 transcribes its Operon 1 co-temporally with the helper phage morphogenetic operon, producing proteins that ultimately lead to formation of smaller sized capsid particles with a head diameter of about 35 micrometers alongside the normal sized phage particles with head diameter of about 48 micrometers (Ruzin *et al.*, 2001). These particles of two distinct sizes can be separated by zonal centrifugation employing sucrose gradients but not by isopycnic centrifugation on CsCl gradients, indicating that they have a similar DNA-protein ratio and that the particles are completely filled. Furthermore, these smaller particles were shown to carry SaPI1 DNA and were capable of transducing SaPI1 (Ruzin *et al.*, 2001). These particles have been shown to consist entirely of helper phage encoded proteins for both SaPI1 (Tallent *et al.*, 2007) and SaPIbov1 (Tormo *et al.*, 2008). Genetic experiments in SaPIbov1 have shown that

inactivation of gp8 or gp9 causes failure to produce SaPIbov1 monomer sized DNA, implying that SaPIbov1 gp8 and gp9 are involved in production of smaller sized capsids. In addition to capsid size redirection, high frequency mobilization of SaPI follows a novel mechanism involving small terminase swapping that leads to specific headful packaging of SaPI DNA into helper phage encoded capsids. This phenomenon of small terminase swapping is a unique feature of SaPI mobilization. Deletion of SaPI *terS* results in a thousand fold decrease in the mobilization frequency of SaPI1 (Ubeda *et al.*, 2009). Helper phage tail assembly occurs independently and is not altered by SaPIs. Finished tails attach to the filled capsid by docking to the portal complex yielding infectious SaPI virions. Expression of 80α late genes leads to production of holin and endolysin that lyse cells and release infectious SaPI and phage virions. These SaPI particles can infect naïve cells and complete their life cycle by re-integrating at their specific chromosomal *att* site.

Staphylococcal temperate phages

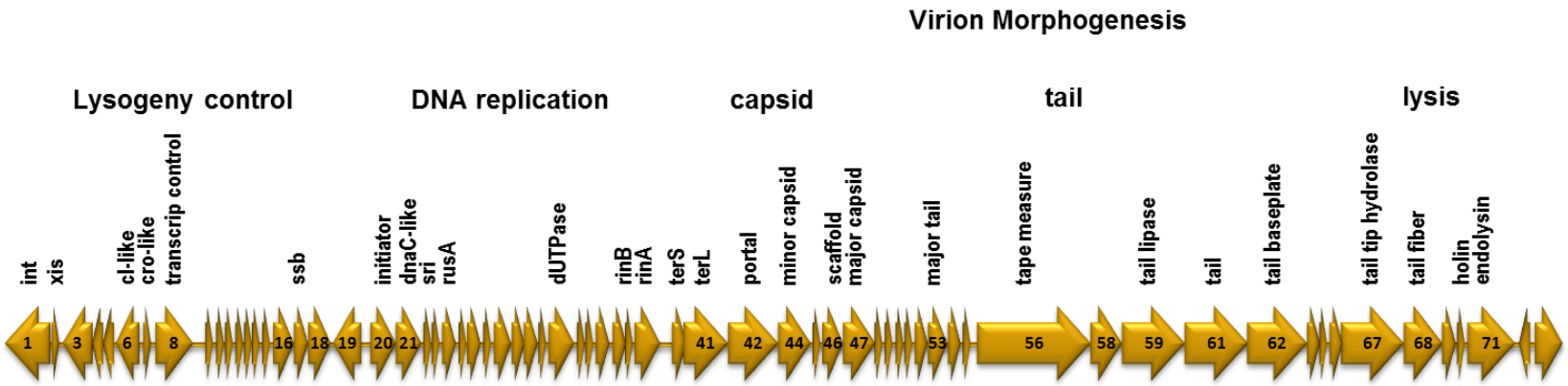
A majority of *S. aureus* phages are temperate with a genome size of approximately 45kb. However, there are also a few virulent phages, including *Myoviridae* such as Twort and Φ29-like *Podoviridae*. All *S. aureus* temperate phages sequenced thus far (Landolo *et al.*, 2002; Vybiral *et al.*, 2003; Kwan *et al.*, 2005; Christie *et al.*, 2010) belong to the family *Siphoviridae* and order *Caudovirales* and have icosahedral heads with non-contractile flexuous tails. These phages fall under two morphotypes, namely morphotype B1 and morphotype B2 (Ackermann, 2001). The majority of *Siphoviridae,* including phage 80α, belong to morphotype B1 and have isometric icosahedral capsids. A few *Siphoviridae* such as HK2, N9, N15, P52 and P87

belong to morphotype B2 and have elongated icosahedral capsids (Hans-Wolfgang Ackermann, 2001). All these phages share a modular genomic organization that is highly mosaic, with genes clustered into functional units that presumably allow temporal expression of polycistronic transcripts. Gene expression in these phages is highly ordered and regulated.

Staphylococcus phage 80α

S. aureus phage 80α (GenBank: DQ517338.1) is a generalized transducing phage that mobilizes several different SaPIs. The length of the 80α genome is 43,864 bp and it contains 73 open reading frames (orfs) of 50 or more codons (Fig 3). Phage 80α was first isolated by Novick (Novick, 1963) during attempts to adapt phage 80 for growth on NTCT8325, and was originally believed to be a recombinant of phages 11 and 80. However recent studies clearly demonstrate that phage 80α is a variant of typing phage 53 that has recombined with phages 11 and 13 (Christie *et al.*, 2010). Phage 80α has nearly 90% nucleotide sequence identity with phage 53. The areas of divergence between phage 80α and phage 53 are accounted for by nearly identical matches from sequences of phages 11 and 13, except for a 2663 bp block of sequence affecting the immunity module that is not found in either phage 11 or 13, but is highly conserved among other staphylococcal *Siphoviridae* (Christie *et al.*, 2010).

Figure 3. Map of the 80α genome.
Shown is a schematic representation of the 80α genome with its modular architecture. Relevant genes are indicated as well as putative functional designations.


Virion Morphogenesis
Lysogeny control
DNA replication
capsid
tail
lysis
int
xis
cI-like
cro-like
transcrip control
ssb
initiator
dnaC-like
sri
rusA
dUTPase
rinB
rinA
terS
terL
portal
minor capsid
scaffold
major capsid
major tail
tape measure
tail lipase
tail
tail baseplate
tail tip hydrolase
tail fiber
holin
endolysin
1
3
6
8
16
18
19
20
21
41
42
44
46
47
53
56
58
59
61
62
67
68
71

80α
43864 bp

Like the other staphylococcal *Siphoviridae*, phage 80α has a distinctly modular genomic architecture with clustering of genes that are temporally co-expressed. The phage 80α genome can be divided into eight different modules, namely (1) phage integration, (2) transcriptional regulation (3) DNA replication (4) packaging (5) head (6) tail (7) tail fibers and (8) lysis, as per the model originally proposed by Botstein (Botstein, 1980).

The phage integration module consists of the phage attachment site (*attP*), phage integrase gene, and phage excisionase gene (*xis*). Phage 80α integrates in an intergenic region between genes *rmpF*, which encodes ribosomal protein L32, and *sirH* which encodes cell wall anchored protein SirH. The integration site corresponds to nucleotide position 1,042,159 in the NCTC 8325 genome (accession number: NC_007795). Phage 80α integrase belongs to the serine recombinase family and is required for integrative and excisive recombination. Excisionases are architectural proteins that bind to DNA and bend it to promote attL×attR recombination. Though phage 80α and phage 53 encoded integrases are identical, their respective excisionases are not. Phage 80α excisionase is identical to Ø11 *xis*. The mechanism of recombination by serine recombinases in staphylococcal *Siphoviridae* is poorly understood and it is not completely known how they promote excision in spite of shuffling their excisionase.

Like that of other temperate phages, the 80α lysogeny module consists of a bidirectional switch region that encodes two divergent transcriptional regulators responsible for the lysis-lysogeny decision. The lytic-lysogeny bistable switch has been extensively studied in phage λ (Dodd *et al.*, 2005). Lysogeny is a stable epigenetic state

during which a master repressor, cI in the case of phage λ, represses promoters P_R and P_L and thereby prevents transcription of lytic genes, including *cro*, while activating P_{RM} to provide constitutive production of cI. Like CI_λ, phage 80α repressor protein consists of two domains. The N-terminal domain has a helix-turn-helix motif consistent with DNA binding activity while the C-terminus has a LexA-like autopeptidase domain found in a large number of phage repressors. The presence of a C-terminal LexA-like domain indicates that, like LexA, cleavage of 80α repressor is mediated by RecA during SOS induction of an 80α prophage. The Cro-like proteins of 80α also contain a helix-turn helix motif, characteristic of DNA binding ability.

80α has an initiator-helicase loader type of replication module (Weigel *et al.*, 2006; Christie *et al.*, 2010) which is preceded by a large orf of unknown function on the complementary strand. *Siphoviridae* of gram positive bacteria with an IL-type (initiator-helicase loader) of replication module encode initiators of λ O-type but helicase loaders of $DnaC_{Eco}$- type (Weigel *et al.*, 2006). The other important characteristics of this type of replication module are that the phage replication origin resides within the initiator gene and the initiator is also located upstream of the helicase loader. Resolvase genes are invariably found downstream of the initiator-helicase loader gene pairs and are mostly *rusA*-type. The phage integrase and repressor genes are always found upstream of the initiator and transcribed in the opposite direction. The putative initiator protein in 80α is encoded by orf 20, which has a highly conserved C-terminal domain also found in other bacteriophages. Within the DNA sequence that codes for the N-terminal region of this protein lies the presumptive origin of replication of this phage. The origin of replication is characterized by direct and inverted repeats and is AT rich. The putative helicase

loader, DnaC, is encoded by orf 21 and is highly conserved among other staphylococcal *Siphoviridae* such as ΦSLT, 80, ROSA, 69, 85, 77, 88 and ΦETA. Downstream of the helicase loader is an open reading frame (orf 22) that encodes Sri. In-frame deletion of *sri* does not affect phage growth but fails to induce SaPI1 (Tallent, 2007; Harwich, 2009; Tormo-Mas *et al.*, 2010). Downstream of *sri* is a gene (orf 23) that has a highly conserved domain but its function is unknown. Orf 24 encodes a RusA type resolvase. RusA type resolvases are homodimeric Holliday junction specific endonucleases that resolve Holliday intermediates by junction cleavage (Sharples *et al.*, 1999). Between the replication and packaging modules are sixteen orfs of which only two encode genes of known function. The dUTPase encoded by orf 32 is a moonlighting protein that not only has dUTPase activity but also acts as an antirepressor of SaPIbov1 (Tormo-Mas *et al.*, 2010). Orf 39 encodes RinA, which has been recently implicated as a positive regulator of transcription of the morphogenetic genes in 80α and Ø11 (Harwich, 2009; Ferrer *et al.*, 2011).

The packaging module consists of genes that encode the terminase small (*terS*) and large subunits (*terL*) and portal proteins (gp42). In most double stranded viruses that package their DNA by the headful mechanism, the small and large terminase subunits assemble to form a holoenzyme, the terminase complex, responsible for recognition of a packaging (*pac*) site on replicating concatemeric phage DNA and making an initial cleavage downstream of this *pac* site (Fujisawa *et al.*, 1997). The DNA bound terminase nucleoprotein docks to the portal to make a packaging motor. The packaging motor translocates the viral genome into the procapsid by a mechanism known as thermal ratcheting until it is condensed to a near liquid crystalline density. The

packaging process is a highly energy intensive process that is driven by ATP hydrolysis and the packaging motors are the most powerful biological motors thus far known. Upon packaging a headful of DNA, the terminase makes a second cut thus separating the DNA concatemer- terminase complex from the filled capsid. This concatemer-terminase complex then docks to another procapsid to initiate a second round of packaging, and so on. DNA packaging is thus a processive process and each round of packaging ensures that a DNA length of more than the viral genome is packaged into each capsid (Johnson *et al.*, 2007).

The capsid morphogenesis module consists of genes encoding minor capsid protein (gp44), scaffold (gp46) and major capsid protein (gp47). Procapsid assembly is presumably initiated by interaction of scaffold with portal proteins. Following this initial nucleation, capsid protein self assembles to form a transient capsid precursor known as procapsid. The procapsid undergoes maturation during the packaging process to form mature capsid. One intriguing feature of SaPI1-phage interaction is capsid size redirection. Helper phage capsid assembly is redirected to form smaller capsids due to presence of SaPI encoded proteins. Capsid morphogenesis and size redirection are discussed in greater detail in later sections.

Tail assembly is a complex process that is as elaborate as the head assembly in *Caudovirales*. In *Podoviridae*, tail assembly occurs by sequential attachment of tail proteins to the filled capsid. In Φ29, for example, the upper collar protein is assembled first followed by lower collar, tail and neck appendages (Camacho *et al.*, 1979). Tail assembly in *Myoviridae* and *Siphoviridae* occurs independently of head assembly and maturation, leading to formation of a “mature” tail that binds to the head via the neck

proteins. The tail assembly in *Siphoviridae* and *Myoviridae* follows a strict order in spite of being an independent process. Tail assembly has been extensively studied in phage T4, and found to be strictly sequential (Kikuchi *et al.*, 1975a; Kikuchi *et al.*, 1975b; Kikuchi *et al.*, 1975c). In *Siphoviridae*, tail assembly starts from the initiator complex that forms the absorption device at the distal end of the tail, followed by polymerization of major tail proteins to form a tubular structure. Binding of terminator proteins to the tubular complex completes the tail assembly (Pell *et al.*, 2009). These preformed tails in turn interact via the terminator proteins with the neck proteins that are associated with the head to form a tailed virus. The length of the tails is determined by the tape measure protein that presumably acts as an internal scaffold for the polymerization of the tube. 3D-reconstructions of cryo-EM images of T4 and SPP1 tails reveal electron density inside their tail tube that could be attributed to either DNA or the tape measure protein (Leiman *et al.*, 2004; Kostyuchenko *et al.*, 2005; Plisson *et al.*, 2007).

The tail module of phage 80α consists of genes encoded by orf 49 to orf 62. The protein encoded by orf49 had been identified as a structural component of the phage and belongs to the family of putative head-tail connector proteins that also include the proteins gp6 and gp15 from bacteriophage HK97 and SPP1, respectively. Orf 50, 51 and 52 encode proteins with unknown functions. The product of orf51 belongs to the family DUF646 and is predicted to be a protein involved in tail completion. The major tail protein encoded by orf 53 is a highly abundant protein that belongs to phage tail 2 super family and is found highly conserved among Gram-positive phages. In-frame deletion of orf 53 is lethal to the phage as well as to SaPI1. There are two small overlapping orfs immediately downstream of the gene encoding major tail protein. These overlapping

orfs contain a predicted -1 translational shift characterized by heptanucleotide “slippery sequences” (Xu *et al.*, 2004; Christie *et al.*, 2010). This type of programmed translational frameshift was first identified in phage λ (Levin *et al.*, 1993) and has been found in many tailed bacteriophages. The tape measure protein, encoded by orf 56, has the largest predicted mass (125.77 KDa) among the 80α proteins. It has been identified as a structural protein of low abundance and is completely absent in the procapsid fraction (Tallent *et al.*, 2007). Orf 58 encodes a protein that belongs to the sipho-tail super family and is a minor component of the tail (Poliakov *et al.*, 2008). The proteins encoded by orfs 59, 61 and 62 are found in both phage and transducing particles and are predicted to be minor tail proteins (Tallent *et al.*, 2007; Christie *et al.*, 2010). The protein encoded by orf 59 has two conserved domains, a DUF1142 domain of unknown function found in phages at the N-terminal end and a SGNH hydrolase domain belonging to a subfamily of lipases and esterases at the C-terminal end (Tallent *et al.*, 2007). The protein encoded by orf62 is essential for infectious phage particles but not required by SaPI1 transducing particles, for reasons not understood. Furthermore, in-frame deletion of the orf62 homolog in a closely related phage, Φ11, yields particles lacking base-plate like structure at the end of tails in both phage and transducing particles (Tormo *et al.*, 2008; Christie *et al.*, 2010).

The tail fiber module of phage 80α consists of orfs 67 and 68. Orf67 encodes a protein that has an N-terminal CHAP domain and a C-terminal lysozyme domain and presumably has peptidoglycan hydrolase activity. Orf 68 encodes a putative tail fiber protein containing a central collagen triple helix motif found in tail fibers of other phages (Christie *et al.*, 2010).

Bacteriophages employ two fundamentally different strategies for bacterial cell lysis (Wang *et al.*, 2000). Most ds DNA phages use a strategy that involves at least two proteins, namely holin and lysin. The holin is a transmembrane protein that creates lesions in the cytoplasmic membrane through which endolysins can pass and gain access to the murein layer. In the other strategy, used by most ssRNA and ssDNA phages such as phage ΦX174, there is no involvement of endolysin. Instead, there is a small trans-membrane protein such as the E-protein in ΦX174, having no sequence similarity to holin, that disrupts the cytoplasmic membrane to cause cell lysis. All phages of the gram-positive bacteria have endolysins with two domains (Diaz *et al.*, 1990; Garcia *et al.*, 1990). The N-terminal domain contains the catalytic activity that may be either an endo-β-N-acetylglucosaminidase or N-acetylmuramidase while the C-terminal cell binding domain (CBD domain) helps it bind to specific carbohydrates found in the cell wall of the host bacteria (Lopez *et al.*, 1992). Timing is of utmost essence in bacterial cell lysis because premature lysis would invariably lead to inefficient yield of phage progeny. The lytic module is expressed late after completion of replication and leads to accumulation of the transmembrane holin protein. Recently it has been shown with help of elegant experiments in *E. coli* that holin builds up to a critical level that triggers the lysis clock (Grundling *et al.*, 2001; White *et al.*, 2011). Phage 80α holin is encoded by orf 70 and is predicted by MEMSAT3 topology analysis (Nugent *et al.*, 2009) to contain two transmembrane helices. Phage 80α endolysin is encoded by orf71 and contains three conserved domains, namely the N-terminal CHAP domain followed by an amidase 3 and a C-terminal SH3b cell wall binding domain. The N-terminal CHAP (cysteine, histidine-dependent amidohydrolases/ peptidases) domain is often found in

association with other domains that cleave peptidoglycan (Bateman *et al.*, 2003). N-acetylmuramoyl-L-alanine amidase or MurNAc-LAA is an autolysin that hydrolyzes the amide bond between N-acetylmuramoyl and L-amino acids in certain cell wall glycopeptides. The C-terminal SH3b cell wall binding domains have been shown to be necessary for accurate cell wall recognition and subsequent staphylolytic activity for some endolysins (Becker *et al.*, 2009).

The SaPI1- helper phage 80α relationship

SaPI1 has a parasitic relationship with helper phage 80α and exploits many of the phage functions for its own propagation. The molecular piracy exhibited by SaPI1 is reminiscent of the exploitation of helper phage P2 by an unrelated satellite phage/ plasmid P4 in *E. coli*. The intricate relationship between P2 and P4 serves as a useful model to explain molecular piracy and has been extensively studied (Inman *et al.*, 1971; Barrett *et al.*, 1973; Six *et al.*, 1973; Six,1975; Shore *et al.*, 1978; Six *et al.*, 1978) and reviewed in detail (Christie *et al.*, 1990; Lindqvist *et al.*, 1993).

The P2-P4 paradigm

In the P2-P4 system, each phage is able to induce the other via transcriptional cross talk. When P2 infects a P4 lysogen, P2 encoded Cox induces P4 by functioning as a transcriptional activator of the P4 P_{LL} promoter and thereby causing it to excise and replicate. When P4 infects a P2 lysogen, P4 encoded ε-protein derepresses P2 to excise and replicate. Furthermore, satellite phage P4 does not encode any structural proteins except an external scaffold, Sid, and relies completely on helper phage P2 structural proteins for its own propagation. Like most dsDNA bacteriophages, P2 capsid

assembly involves copolymerization of major capsid proteins (gpN) and portal proteins (gpQ) in the presence of internal scaffolding protein (gpO). Size determination of the P4 capsid is carried out by a P4 encoded protein, Sid, that acts transiently as an external scaffold (Marvik *et al.*, 1995) during capsid assembly and favors the formation of smaller capsids (T=4) to the larger ones (T=7). These smaller capsids have roughly 1/3rd the volume of P2 capsids, commensurate with the smaller size of the P4 genome (Dokland *et al.*, 1992). Additionally, both phages cause transactivation of late gene expression from the other genome. P4 encoded δ transactivates expression of P2 morphogenetic operons, along with its own late gene expression. Similarly, P2 encoded Ogr is a transcriptional activator of both P2 and P4 late gene expression. Thus P4 efficiently exploits P2 functions at various levels for its own propagation.

The molecular piracy exhibited by P4 is a useful model to explain, albeit not completely, SaPI-helper phage interaction. Like P4, SaPIs lack structural genes and rely completely on helper phage functions for their own encapsidation at the expense of phage propagation. Normally, 80α has a burst size of about 600 infectious particles but in presence of SaPI1 is reduced by two orders of magnitude to about six infectious particles per cell (Ruzin *et al.*, 2001). This phenomenon is termed SaPI interference and its effect on helper phage propagation is so profound that plaque formation is completely blocked if phage is plated on a lawn of a SaPI1-positive strain. There are at least three levels of interaction between SaPI1 and its helper phage. The first interaction is at the transcriptional level. SaPI1 excision and replication are helper phage dependent and rely on 80α encoded Sri for derepression of SaPI1 early rightward gene expression. When 80α infects a SaPI containing strain or following induction of a strain

containing both 80α and SaPI1, SaPI1 is derepressed and commits to the lytic decision. However, unlike the P2-P4 system, transduction of SaPI1 into an 80α lysogen does not invoke lytic mode in 80α because there is no reciprocal derepression. The second level of interaction involves small terminase swapping, a unique feature that is not found in the P2-P4 system. SaPI1 encoded TerS efficiently interacts with 80α encoded TerL to encapsidate SaPI1 DNA resulting in a transduction frequency at least three orders of magnitude above the level found in generalized transduction. In-frame deletion of SaPI1 *terS* drastically decreases the yield of transducing particles. The third level of interaction between SaPI1 and 80α involves capsid assembly. SaPI1 encoded proteins interact with helper phage major capsid proteins to redirect capsid size and form smaller capsids which are one-third the size of 80α capsids. These small capsids are able to encapsidate a SaPI1 monomer but not a complete phage 80α genome.

This study was designed to understand the mechanism of helper phage capsid size redirection. Several approaches were used. In order to elucidate the morphology of 80α and SaPI1 capsids and procapsids, we constructed mutants to allow efficient isolation of capsid precursors for structural analysis. In collaboration with Dr. Terje Dokland's lab at UAB, we conducted structural studies that involved isolation of these particles and subjected them to cryo-electron microscopy and 3D-reconstruction. Modeling studies were performed to understand how the capsid subunits interacted with each other. SDS-PAGE and mass spectrometric studies of capsid precursors were used to identify SaPI1 proteins that could be involved in capsid size redirection. Targeted mutagenesis and co-expression studies were carried out to establish the roles of

individual proteins in capsid assembly and size redirection and to understand the phenomenon of interference.

Chapter 2

Materials and Methods

Bacterial culture methods

Bacterial strains that were used in this study are listed in Table 1. *S. aureus* was grown overnight on phage agar (0.3% (wt/vol) Casamino acids, 0.3% (wt/vol) yeast extract, 100 mM NaCl, 1.5% agar (wt/vol), 0.5 mM CaCl2, pH 7.8)(Novick,1991) or on Tryptic Soy Agar (TSA) (Remel, Lenexa, KS) at an incubation temperature of 30ºC. *S. aureus* strains that contained SaPI1 *tst::tetM* were plated on GL agar (Novick,1991) (0.3% (wt/vol) Casamino acids, 0.3% (wt/vol) Yeast extract, 100 mM NaCl, 0.33% (vol/vol) Sodium lactate syrup, 25% (vol/vol) glycerol, 1.5% (wt/vol) agar, 0.17 mM sodium citrate, pH 7.8) supplemented with 5 µg/ml tetracycline. *S. aureus* strains were cultured in either Tryptic Soy Broth (TSB) (Remel, Lenexa, KS), Brain Heart Infusion (BHI) (Remel, Lenexa, KS), or CY-GL broth (Novick, 1991) and grown at 30ºC unless otherwise indicated. SA178RI derivatives were grown overnight at 37ºC either in BHI or Antibiotic medium 3 supplemented with appropriate antibiotics. *E. coli* strains were cultured in Luria Bertani (LB) broth (Difco, Franklin Lakes, NJ) or on LB Agar plates supplemented with ampicillin (100 µg/ml), as required and grown at 37ºC.

Allelic exchange using derivatives of the shuttle vector pMAD required media

Table 1. Bacterial strains and plasmids used in this study

***S aureus* strain**	**Description**	**Reference or source**
RN450	NCTC8325 cured of Φ11, Φ12 and Φ13	Novick, 1967
RN4220	Restriction defective derivative of RN450	de Azavedo *et al.*, 1985
SA178RI	CYL316 containing T7 RNA polymerase, Tet^R	D'Elia *et al.*, 2006
RN10616	RN4220 (80α)	Ubeda *et al.*, 2009
RN10628	RN4220 (80α) SaPI1 *tst::tetM*	Ubeda *et al.*, 2009
ST24	RN4220 (80α Δ*terS*)	This study
ST37	RN4220 (80α) SaPI1 *tst::tetM* ΔterS	This study
ST51	RN4220 (80α) Δ*orf46* SaPI1 *tst::tetM*	This study
ST63	RN4220 (80α) Δ*terS* SaPI1 tst::tetM ΔterS	This study
ST64	RN4220 (80α) *Δorf44*	This study
ST65	RN4220 (80α) *Δorf44* SaPI1 *tst::tetM*	This study
ST66	SA178RI pPD18	This study
ST70	SA178RI pPD20	This study
ST71	SA178RI pPD21	This study
ST72	SA178RI pPD22	This study
ST73	SA178RI pG164	This study
ST82	RN4220 (80α::*SaPI1 orf6+7)*	This study
ST83	RN4220 (80α *ΔterS Δorf53)* SaPI1 *tst::tetM*	This study
ST91	RN4220 (80α *Δorf46)*	This study
ST92	RN4220 (80α *Δorf44 Δorf53)* SaPI1 *tst::tetM*	This study
ST97	RN4220 (80α::*SaPI1 orf7)*	This study
ST98	RN4220 (80α) SaPI1 *tst::tetM Δorf7*	This study
ST99	RN4220 (80α::*SaPI1 orf6)*	This study
ST100	RN4220 (80α) SaPI1 *tst::tetM Δorf6*	This study
ST112	SA178RI pPD44	This study
ST113	SA178RI pPD45	This study
ST114	SA178RI pPD46	This study
ST115	SA178RI pPD47	This study
ST116	SA178RI pPD49	This study
ST117	SA178RI pPD50	This study
ST118	SA178RI pPD51	This study
ST119	SA178RI pPD52	This study
***E. coli* strains**		
DH5α	*E. coli* F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K^- m_K^+), λ–*	Invitrogen
BLRDE3	*E. coli F- ompT* $hsdS_B$*(*r_B^- m_B^- *) gal dcm(DE3) Δ(srl-recA)306::Tn10 (*Tet^R*)*	Novagen
Plasmids for over-expression of proteins		
pET21a	Plasmid containing T7 promoter for expressing proteins in *E. coli*	Novagen
pG164	*S. aureus* shuttle plasmid containing T7 promoter	D'Elia *et al.*, 2006
pPD1	pET21a derivative with 80α gp47	This study
pPD2	pET21a derivative with 80α gp47 and gp46	This study
pPD18	pG164 derivative with 80α gp47	This study

pPD20	pG164 derivative with 80α gp47.	This study
pPD21	pG164 derivative with 80α gp46 and gp47	This study
pPD22	pG164 derivative with 80α gp42 to gp47	This study
pPD44	pG164 derivative with 80α gp47 and SaPI1 gp6	This study
pPD45	pG164 derivative with 80α gp47 and SaPI1 gp7	This study
pPD46	pG164 derivative with 80α gp47 and SaPI1 gp6 and gp7	This study
pPD47	pG164 derivative with 80α gp47 and SaPI1 gp5, gp6 and gp7	This study
pPD51	pG164 derivative with 80α gp46, gp47 and SaPI1 gp6 and gp7	This study
pPD52	pG164 derivative with 80α gp46, gp47 and SaPI1 gp5, gp6 and gp7	This study
Plasmids for deletion		
pMAD	Shuttle plasmid used for allelic exchange in Gram-positive bacteria	Arnaud *et al*., 2004
pPD8	pMAD derivative for deletion of 80α *orf 46*	This study
pPD11	pMAD derivative for deletion of SaPI1 *terS*	This study
pPD14	pMAD derivative for deletion of 80α *orf22*	This study
pPD17	pMAD derivative for deletion of 80α *orf44*	This study
pPD34	pMAD derivative for deletion of SaPI1 *orf7*	This study
pPD35	pMAD derivative for deletion of SaPI1 *orf6*	This study
pPD27	pMAD derivative for deletion of 80α *orf 53*	This study
Plasmids for inserting mutant alleles		
pPD36	pMAD derivative for inserting SaPI1 *orf6* in 80α	This study
pPD37	pMAD derivative for inserting SaPI1 *orf7* in 80α	This study
pPD28	pMAD derivative for inserting SaPI1 *orf6 and7* in 80α	This study

supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal) (200 μg/ml) (American Bioanalytical, Natick, MA) and erythromycin (5 μg/ml) for staphylococcal strains or ampicillin (100 μg/ml) for *E. coli*. Studies involving overexpression of proteins in *S. aureus* strain SA178RI containing pG164 derivatives were grown overnight in media containing 15 μg/ ml chloramphenicol and 5 μg/ml tetracycline.

Phage propagation

Freshly prepared BHI or CY-GL media were inoculated, at a ratio of 1:100, with previously grown overnight broth culture of a phage-sensitive *S. aureus* strain, RN4220, RN450, or their derivatives, and incubated at 30ºC on an orbital shaker at 200 rpm. Cell growth was monitored using either a Klett-Sommerson colorimeter or a spectrophotometer until it reached an optical density (OD) of 0.6 ($\lambda_{absorbance}$ = 550nm) (Klett = 50). This optical density corresponds to the mid-exponential phase of growth, with a cell concentration of about 3 X 10^8 cfu. At this point an equal volume of SA Phage Buffer (Novick,1991)(1 mM MgSO4, 4 mM CaCl2, 0.05 M Tris-HCl pH 7.8, 100 mM NaCl, 0.1% gelatin) was added to the culture, followed by addition of bacteriophage 80α at a multiplicity of infection (m.o.i.) equal to 0.1. The infected culture was incubated at 30ºC on an orbital shaker at 100 rpm and monitored for lysis until the optical density stopped decreasing.

Alternatively, bacteriophage 80α was obtained by induction of lysogenic staphylococcal strains. Freshly prepared BHI medium was inoculated 1:100 with overnight cultures of lysogenic strains and incubated at 30^{o}C on an orbital shaker at 200 rpm until Klett=50. Lysogeny was terminated either by addition of mitomycin C (final

concentration=2 µg/ml) or ciprofloxacin (final concentration= 1 µg/ml) as indicated, or by UV irradiation. Mitomycin C is a potent DNA cross linker while ciprofloxacin acts by trapping DNA gyrase on DNA and thus blocking the replication fork movement. For UV induction, the cells were spun down by centrifugation for 15 minutes at 5000 rpm at 4ºC, resuspended to the original volume in SA phage buffer, exposed to UV light (intensity= 70 Joules.cm^{-2}) for 20 seconds or as indicated, diluted 1:1 with BHI broth, and incubated at 30ºC, 200 rpm until lysis. These lysates were filtered through 0.45 µm syringe filters (Millipore, Billerica, MA) and stored at 4ºC for further analysis.

Phage concentration was determined by plating tenfold serial dilutions of lysate on a sensitive *S. aureus* indicator strain (RN450 or RN4220) using the double layer method. Phage were serially diluted in SA phage buffer, and 100 µl aliquots of each phage dilution were added to 100 µl of an overnight culture of indicator cells, mixed with 2 ml of SA top agar (Novick, 1991) (0.3% (wt/vol) Casamino acids, 0.3% (wt/vol) Yeast extract, 100 mM NaCl, 0.5% agar (wt/vol), 0.5 mM CaCl2, pH 7.8) and quickly poured onto the surface of a phage agar plate. The plates were then initially incubated upright for about 30 minutes until the top agar had hardened and then inverted and left overnight in the incubator at 30ºC. Plaques were counted and phage titer was determined using the following equation:

Pfu/ml = number of plaques X 10 / dilution

The transducing titer of lysates prepared from SaPI1 *tst*::*tet*M strains was determined by using ten-fold serial dilutions of lysates. 100µl of these dilutions were mixed with 100µl of indicator cells as described above. The mixture was spread onto the surface of GL agar plates supplemented with tetracycline (5 µg/ml). Colonies were

counted after a 48-hour incubation period at 30ºC. Transducing units were determined using the equation:

TU/ml = number colonies X 10 / dilution

PEG precipitation

Cellular debris was removed from phage 80α and SaPI1 lysates by centrifugation at 7000 rpm for 20 minutes at 4ºC in a Sorvall GS-3 rotor. Polyethylene glycol 8000 (PEG) (0.1% wt/vol) and NaCl (0.5 M) were slowly dissolved, with gentle stirring, in the supernatant and the mixture was incubated overnight at 4ºC. The lysate was centrifuged for 20 minutes at 7000 rpm and 4ºC to precipitate phage containing PEG. The supernatant was decanted and the PEG precipitate was removed from the sides of the bottle using a cell scraper or a rubber policeman and resuspended in SA phage buffer (9 ml per liter of lysate). The resuspended PEG precipitate was collected in a sterile Corex tube and allowed to incubate overnight at 4ºC, followed by centrifugation for 10 minutes at 10,000 rpm at 4ºC in a Sorvall SS-34 rotor. The supernatant was transferred to a sterile tube and the remaining pellet was again resuspended in SA phage buffer (3 ml) to extract the remaining phage and centrifuged a second time under similar conditions. The re-extracted supernatant and the first supernatant were pooled together.

Phage purification

The phage and SaPI1 particles extracted from the PEG precipitate were further purified by centrifugation in a cesium chloride (CsCl) step gradient. Cesium chloride solutions of four different densities corresponding to ρ= 1.3 g/ml, 1.4g/ml, 1.5g/ml and 1.6g/ml were prepared in SA phage buffer as follows: Masses of 10.1 g, 13.47g 16.87g and 20.2g CsCl were dissolved in SA phage buffer to a final volume of 25ml in a

volumetric flask. A step gradient was prepared in a clear centrifuge tube (Beckman, Fullerton, CA) 1 x 3 1/2 inches as follows: The bottom layer was laid with 2ml (ρ=1.6g/ml) of the CsCl solution. 3ml CsCl solution (ρ=1.5g/ml) was carefully laid on top of the first layer with the help of a Pasteur pipette followed by a 4ml (ρ=1.4g/ml) layer and finally a 4ml (ρ=1.3g/ml) layer. The pooled phage supernatant was carefully added on top. The centrifuge tubes containing the CsCl gradients were centrifuged in a Beckman SW28 rotor for three hours at 15ºC 24,000 rpm. The rotor was allowed to stop without braking at the end of the spin. The bluish phage band, visible at the 1.4-1.5 interface, was collected by puncturing the tube at the side with a syringe and removing it. The banded phage sample was dialyzed at 4ºC with four exchanges of SA phage buffer using a Slide-A-Lyzer dialysis cassette (Pierce Biotechnology, Rockford, IL) with a 10,000 molecular weight cut off (MWCO).

Production and purification of capsids and procapsids for EM analysis

80α and SaPI1 procapsids were produced for EM and MS analysis in Dr Dokland's lab in the following way. The 80α Δ*terS* lysogen ST24 was used as the source of 80α procapsids, while SaPI1 procapsids were isolated from the SaPI1Δ*terS*-containing 80α lysogen ST37 and SaPI1Δ*terS*-containing 80α Δ*terS* lysogen ST64. Strains were grown at 32 C in CY broth as previously described, and the 80α prophages were induced by addition of 2 mg/L mitomycin C (Sigma) at OD_{540} = 0.4–0.5, with lysis occurring 3 h post induction. The lysates were clarified by centrifugation at 5400×*g* for 20 min. Procapsids were precipitated with 10% (wt/vol) PEG 6000 and 0.5 M NaCl overnight at 4 °C, followed by centrifugation at 5400×*g* for 20 min. The pellets were

resuspended in phage buffer (50 mM Tris pH 7.8, 100 mM NaCl, 1 mM $MgSO_4$ and 4 mM $CaCl_2$), and 0.30–0.50g of CsCl was added per milliliter of suspension. In some cases, half a volume of chloroform was added to the PEG pellet to facilitate resuspension, resulting in a cleaner sample. No difference was observed in the structure or protein composition between treated and untreated samples. The resulting solution was centrifuged at 339,000×*g* for 20 h in a Beckman NVT90 rotor. The procapsid-containing bands from the CsCl gradients were collected and dialyzed against dialysis buffer (20mM Tris pH 7.8, 50mM NaCl, 1mM $MgCl_2$ and 2mM $CaCl_2$) for further analysis.

To separate the procapsids from tails, the dialyzed CsCl bands were loaded onto 10–40% sucrose gradients in phage buffer and centrifuged for 2h at 110,000×*g* in a Beckman SW41 rotor. Twelve fractions were collected manually from the sucrose gradients and analyzed by SDS-PAGE. Fractions containing predominantly procapsids and procapsids with attached tails were pelleted by centrifugation at 110,000×*g* for 1h. The pellet was resuspended in dialysis buffer and used for EM and MS experiments.

Electron microscopy

Samples for negative stain were prepared by applying 3 µl of procapsid suspension to glow-discharged carbon-only grids (Electron Microscopy Sciences), washed two times with dialysis buffer and stained with 1% uranyl acetate. Cryo-EM was performed by standard methods: 3µl of sample was applied to C-flat holey film (Electron Microscopy Sciences), blotted briefly before plunging into liquid ethane and transferred to a Gatan cryo-sample holder. All samples were observed in an FEI Tecnai F20

electron microscope operated at 200 kV, and images were captured on a 4000 × 4000 Gatan Ultrascan CCD camera or on Kodak SO-163 film at magnifications from 38,000× to 81,200×. Electron microscopy was performed at the HRIF Electron Microscopy Core, University of Alabama, Birmingham.

Protein identification

Coomassie-blue-stained SDS-PAGE gels of 80α and SaPI1 procapsids were cut into 10–12 strips. The strips were destained in 60% methanol and 0.1% trifluoroacetic acid (TFA) and dried with pure acetonitrile. The acetonitrile was removed by evaporation in a Speedvac centrifugal evaporator, and protein digestion was performed by addition of 10 µg/ml Trypsin Gold (Promega) solution in 100 mM ammonium bicarbonate and incubation for 8 h at 37 °C. After extraction with 10% acetonitrile, the peptides were loaded on a 100-nm × 10-cm capillary column in-house packed with C18 Monitor 100 A-spherical silica beads and eluted by a 1-h gradient of 10–100% acetonitrile and 0.1% TFA. Mass spectrometric analysis was performed on an LTQ XL (Thermo Finnigan) spectrometer in the Mass Spectrometry/ Proteomics shared facility, UAB Birmingham. The search for matching peptide sequences was performed using the SEQUEST search engine with the UniProt database, including staphylococcal and phage entries. Only peptides with a probability of >0.99 were considered.

Measurement of full-length protein masses

Measurement of the whole masses of the most abundant proteins composing 80α and SaPI1 procapsids was performed as described by Poliakov *et al* (Poliakov *et al.*, 2008). Briefly, purified procapsids were disrupted in 6M urea and 0.1% TFA, loaded

on a C4 microtrap reverse-phase column and eluted with a gradient of 0–50% isopropanol and 0–50% acetonitrile in 0.1% formic acid. Spectra were acquired on a time-of-flight electrospray mass spectrometer (LCT, Micromass).

Confirmation of cleavage sites of scaffolding and major capsid proteins was achieved by tandem MS of N- and C-terminal peptides on matrix-assisted laser desorption/ionization TOF/TOF tandem mass spectrometer (Ultraflex III; Bruker Daltonics). In-gel digests of 80α scaffolding and major capsid proteins were spotted on a Bruker 384 spot plate and allowed to dry. Equal volumes of 5 mg/ml α-cyano-hydroxycinnamic acid in 60% acetonitrile and 0.1% TFA were spotted on top of the digests and allowed to dry. Peptides matching expected C- or N-terminal peptides of truncated proteins to <0.02Da were fragmented by laser-induced dissociation. Fragment assignment was performed with BioTools software (Bruker Daltonics) using a 0.5-Da fragment mass accuracy cutoff.

DNA manipulations

Restriction endonucleases, T4 DNA ligase, Antarctic phosphatase, PNK kinase, specific buffers and BSA used for DNA manipulation were purchased from New England Biolabs (Ipswich, MA) and used as recommended by the manufacturer. DNA was extracted from agarose gels using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) as described by the manufacturer. PCR products were purified using the QIAquick PCR purification kit (Qiagen, Valencia, CA) as recommended by the manufacturer. Plasmid DNA was isolated from individual *E. coli* transformants that were

grown overnight in 3 ml of LB broth with ampicillin, using a QIAprep Spin Miniprep kit (Qiagen, Valencia, CA,) as described by the manufacturer.

Agarose gels

Agarose gels were prepared by dissolving agarose in 1X TAE buffer at 100ºC (Fisher Scientific, Pittsburgh, PA). Ethidium bromide was added to 0.1 µg/ml prior to pouring the gel. Generally, PCR products were resolved on 1% agarose gels, plasmids and gDNA on 0.7% agarose gels. 5X DNA loading dye (40% sucrose, 0.25% bromophenol blue and 0.25% Xylene cyanol) was added to the DNA solution in the ratio of 1:4 so as to reach 1X dye concentration. DNA was loaded on agarose gels and subjected to electrophoresis at about 5 volts/cm until the bromophenol blue had migrated roughly 3/4 of the gel length and then visualized, under UV light. DNA size was measured against the Hyperladder (BioLine, Taunton MA) series of DNA MW markers.

Plasmid screening

A quick-check method (Akada,1994) was used to screen colonies for plasmids with inserts. Each colony was resuspended in 100 µl of sterile water, to which 50 µl of TE saturated phenol: chloroform: (1:1) plus 10 µl loading dye (0.25% bromophenol blue and 40% glycerol) were added. The tubes were vortexed for 10 seconds and centrifuged for 3 minutes at 13,000 rpm. An aliquot of (10-15 µl) from each sample was loaded onto a 0.7% agarose/TAE gel and compared to a super-coiled DNA ladder (Invitrogen, Carlsbad, CA).

Polymerase Chain Reaction (PCR)

PCR reactions were performed in a T-Gradient thermocycler (Whatman Biometra, Goettingen, Germany). Oligonucleotides used in this study are listed in Table 2. These oligonucleotides were designed from 80α and SaPI1 sequences (GenBank DQ517338 and U93688, respectively) and were purchased from Integrated DNA Technologies (Coralville, IA). Oligonucleotides were reconstituted in HPLC grade water (Mallinckrodt Baker, Phillipsburg, NJ) to 1mM stock solution. Working solutions of oligonucleotides were prepared by diluting these stock solutions to 10 µM concentration. Standard PCR mixtures were set up on ice and prepared as follows: 1X Standard *Taq* Reaction Buffer, DNA template (2-50 pg plasmid or 50-500 ng genomic template), oligonucleotides at 0.5 µM final concentration, 200 µM dNTPs each (Invitrogen, Carlsbad, CA) and 1.25 units of *Taq* DNA Polymerase (NEB, Ipswich, MA) in a total reaction volume of 50 µl. PCR products used for sequencing and cloning were amplified with high fidelity Phusion DNA Polymerase (Finnzymes, Oy -Keilaranta 16 A, 02150 Espoo, Finland.*)* PCR reactions were performed in MQ water with the following components: 1X DNA Phusion buffer, 0.2 mM each dNTP, 0.5 µM each forward and reverse oligonucleotides, DNA template 1 pg–10 ng per 50 µl reaction and DNA Phusion Polymerase (0.04 U/µl final concentration). Cycling conditions for PCR were as follows: Initial denaturation was performed at 98ºC for 30 seconds followed by 29 repetitive cycles of the following operations: denaturation at 98ºC for 10 seconds, primer annealing at hybridization temperature (TM + 3) for 15 seconds and primer extension at 72ºC for 20 seconds per kb of extension. A final extension at 72ºC for 5 minutes was performed after the final cycle and the reactions were chilled to 4ºC.

Preparation of electrocompetent cells

Electrocompetent *E. coli* cells (Miller *et al.*, 1995), (Seidman *et al.*, 2001) were prepared from cultures grown in 250 ml LB (Difco, Franklin Lakes, NJ) at 37°C with shaking at 300 rpm to an OD600 of approximately 0.5–0.7. Cells were chilled on ice for ~20 min. For all subsequent steps, the cells were kept as close to 0°C as possible. All containers were chilled in ice before adding cells. Cells were harvested and spun at 4000 x g for 15 minutes at 4°C. The supernatant was carefully discarded and the pellet gently resuspended in 250 ml sterilized ice-cold 10% glycerol. The resuspended cells were again centrifuged at 4000 x g for 15 minutes at 4°C. The supernatant was again discarded and cells were resuspended in 125 ml of ice-cold 10% glycerol and centrifuged again at 4000 x g for 15 minutes at 4°C. An additional repetition of this step ensured maximum removal of salt and greatly helped to avoid the problem of arcing during electroporation. The pellet was resuspended in 10 ml ice-cold 10% glycerol and transferred to a 30 ml sterile Oakridge tube. The resuspended cells were centrifuged at 4000 x g for 15 minutes at 4°C. The supernatant was carefully poured off and the cells were suspended in a final volume of 1ml of ice-cold 10% glycerol yielding a cell concentration of about 1–3 x 10^{10} cells/ml. This suspension was distributed in aliquots of 50 µl and stored at -70 °C.

Electrocompetent *S. aureus* cells were prepared from cultures grown in BHI at 30°C with shaking at 250 rpm, to ~2 x 10^{8} cells/ml. The cells were then chilled in an ice water bath for 15 min to arrest growth and harvested by centrifugation at 12,000 x g for 15 min at 4 °C. The supernatant was carefully removed and the cell pellet was suspended in 500 ml of sterile, ice-cold water and centrifuged again at 12,000 x g for 15 min at 4 °C; The cells were washed two more times in 500 ml of sterile, ice-cold water.

The cell pellet was suspended in 25 ml of sterile, ice-cold 10% glycerol and transferred to a 30 ml sterile Oakridge tube. The cells were pelleted by centrifugation at 12,000 x g for 15 min at 4°C and resuspended in 2ml final volume of 10% glycerol to a final concentration of about 1 x 10^{10} cells/ml. Cells were distributed in 50 µl aliquots into sterile 1.5 ml microfuge tubes and stored at -70 °C.

Transformation of *E. coli* and *S. aureus*

Ligation reactions (2 µl), or 0.5 µl of purified plasmid DNA, were added to 50 µl of electrocompetent *E. coli* cells previously thawed on ice. This mixture was transferred to a cold electroporation cuvette (Bio-Rad, Hercules, CA) with a gap length of 0.1 cm and pulsed one time using the MicroPulser (Bio-Rad, Hercules, CA) pre-set *E. coli* setting Ec1 (1.8 kV). 1 ml of SOC media (2% (wt/vol) tryptone, 0.5% (wt/vol) yeast extract, 85.5 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 20 mM glucose) warmed to 37ºC was immediately added to the cuvette and cells were gently but quickly resuspended using a 1000 µl pipette. The resuspended cells were transferred to a 2ml microfuge tube and incubated for one hour at 37°C with shaking. Following this recovery period the cells were plated on LB plates with appropriate antibiotic selection.

Plasmid DNA isolated as described above was used to transform *S. aureus*. Plasmid DNA (2 µl) was added to 50 µl aliquots of electrocompetent *S. aureus* and incubated on ice for 20 minutes. The mixture was then transferred to an electroporation cuvette with a gap length of 0.1 cm and pulsed one time using the MicroPulser (Bio-Rad, Hercules, CA) pre-set *S. aureus* setting Sta (1.8 kV, 2.5 msec, 25 µF). Following electroporation, one ml of Brain Heart Infusion (BHI) broth (Remel, Lenexa, KS) was added immediately to the cuvette and this mixture was then transferred to a glass

culture tube. The antibiotic to be used for selection was added at a subinhibitory concentration (erythromycin = 0.05 µg/ml,) and the cells were incubated for 1.5–2 hours with shaking at 30°C. Aliquots (100 µl and 200 µl) were then spread on TSA plates supplemented with appropriate antibiotics and incubated at 30°C for 48 hours.

Creation of deletion and insertion mutants by allelic exchange

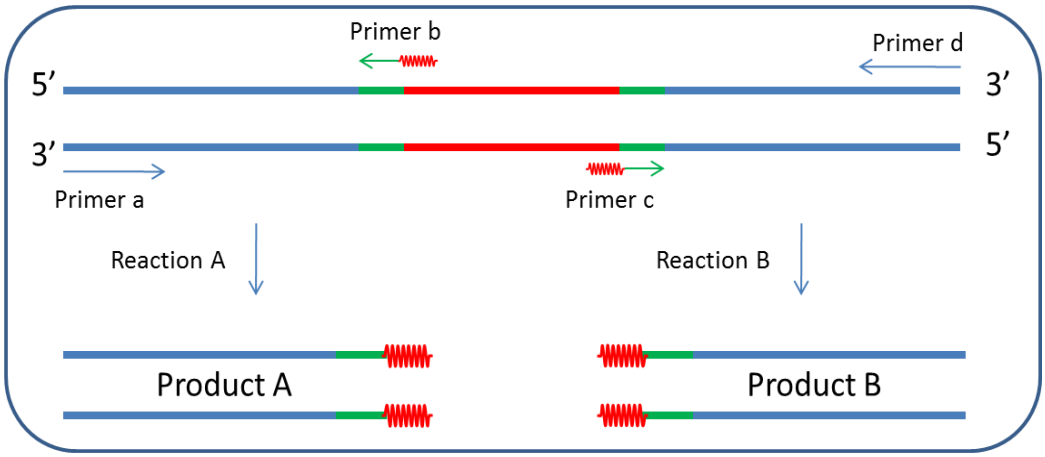
The strategy used for creating deletion and insertion mutants can be divided into the following steps: (1) Creation of the mutant allele using the technique of gene splicing by overlap extension (gene SOEing); (2) Cloning the mutant allele into appropriate restriction sites of shuttle plasmid pMAD; (3) Transforming *E. coli* with the pMAD derivative; (4) Isolation of plasmid from the *E. coli* strain and transforming appropriate *S. aureus* strain and (5) Allelic exchange.

Mutant alleles for deletion

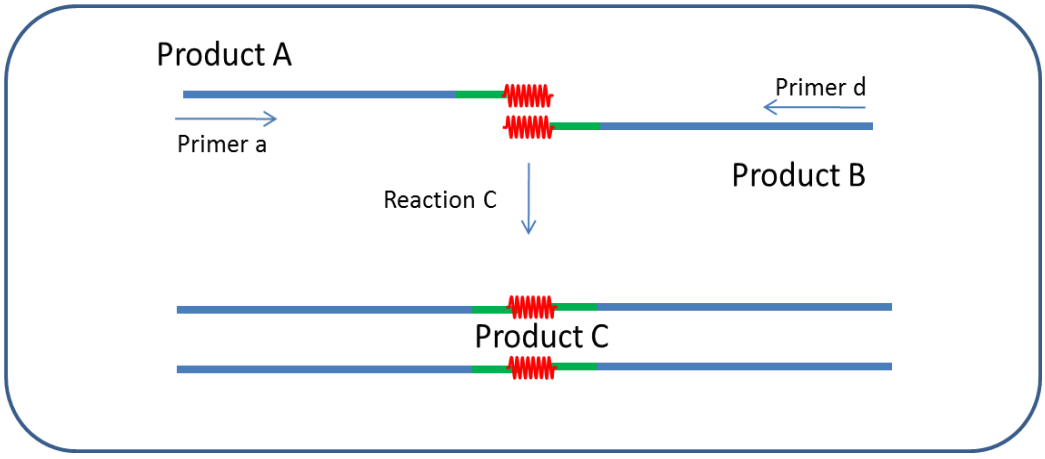
Mutant alleles were generated using the strategy described by Horton (Horton,1995) (Horton *et al.*, 1990). Four different oligonucleotides (Fig 4), a, b, c and d were designed in such a way that primers a and d hybridized at least 400 base pairs away upstream and downstream of flanking regions of the targeted gene. Primers a and d are called the flanking primers and have appropriate restriction sites at their 5’ ends for subsequent cloning. Oligonucleotides b and c, also known as the SOEing primers, complement each other at their 5’ ends and the Tm of this complementary sequence is not less than the Tm of the flanking primers a and d. Also, the Tm of the rest of the 3’ region of the SOEing primers b and c is comparable to the Tms of the flanking primers. PCR reactions A and B were performed with primer pairs a, b and c, d to obtain DNA

Figure 4. Schematic representation of gene SOEing reaction for generating mutant alleles with in-frame deletion

A


Primer b
Primer d
5’
3’
3’
5’
Primer a
Primer c
Reaction A
Reaction B
Product A
Product B

B


Product A
Primer d
Primer a
Product B
Reaction C
Product C

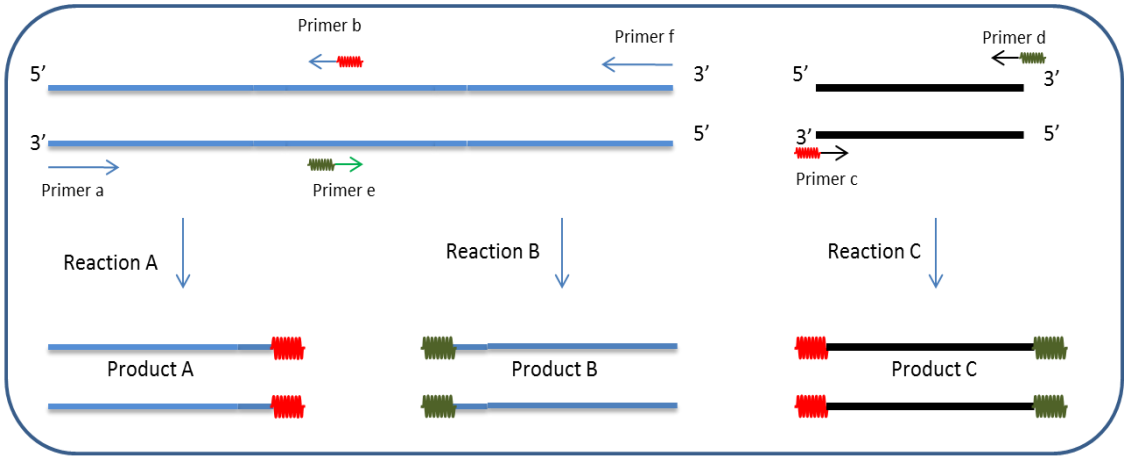
fragments A and B. These DNA fragments were subjected to electrophoresis on agarose gels and the DNA from appropriate bands was extracted using a gel extraction kit (Qiagen, Valencia, CA).The purified DNA was quantified, mixed in an equimolar ratio and used as template for PCR reaction C with flanking primers a and d. The DNA fragment C was gel extracted, digested with appropriate restriction endonucleases and cloned into corresponding restriction sites of the shuttle plasmid pMAD.

Mutant alleles for insertion

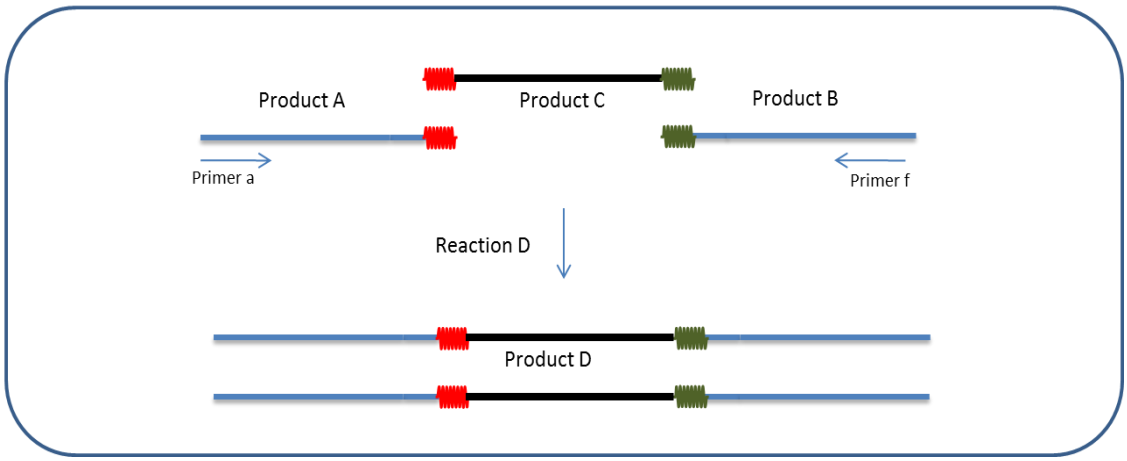
Mutant alleles for insertion were generated by using the same strategy as described above except that six different oligonucleotides, a, b, c, d, e and f were used (Fig 5). The oligonucleotides were designed in such a way that primers a and f hybridized at least 400 base pairs upstream and downstream of flanking regions of the targeted gene. Primers a and f are the flanking primers and have appropriate restriction sites at their 5' end. Oligonucleotide pairs b and c, d and e complement each other at their 5' ends and the Tm of the complementary part of their sequence is not less than the Tm of the flanking primers a and f. PCR reactions A, B and C were performed with primer pairs a-b, e-f and c-d to obtain DNA fragments A, B and C. These DNA fragments were subjected to electrophoresis on agarose gels and the DNA from appropriate bands was extracted using a gel extraction kit. The purified DNA was quantified, mixed in an equimolar ratio and used as template for PCR reaction D with flanking primers a and f. The DNA fragment D was gel extracted, digested with appropriate restriction endonuclease(s) and cloned into the shuttle plasmid pMAD.

Figure 5. Schematic representation of gene SOEing reaction for generating mutant alleles with insertion

A


Primer b
Primer f
Primer d
5′
3′
5′
3′
3′
5′
3′
5′
Primer a
Primer e
Primer c
Reaction A
Reaction B
Reaction C
Product A
Product B
Product C

B


Product A
Product C
Product B
Primer a
Primer f
Reaction D
Product D

Electrocompetent DH5α was transformed with the ligation mixture and plated on LB agar plates supplemented with 100µg/ml ampicillin. The plates were incubated overnight and single colonies were picked and restreaked on LB agar plates supplemented with 100 µg/ml ampicillin.

Plasmid pPD11 was constructed as follows: Flanking DNA fragments of ~1kb each were amplified by polymerase chain reaction (PCR) and ligated together. A region from orf 6 to the beginning of SaPI1 *terS* was amplified from SaPI1 *tst*::*tetM* with primers SMT53 and SMT54 (Table 2), and a region from *tetM* to the end of SaPI1 *terS* was amplified with primers SMT51 and SMT52. The PCR products were digested with HindIII, ligated together and used as template for amplification with SMT51 and SMT54. After purification, the resulting PCR product was digested with NcoI and ligated with NcoI-digested pMAD. The resulting plasmid was introduced into *E. coli* DH5α by electroporation.

Colonies were screened for plasmids with inserts using the quick check method. Plasmids were isolated and verified by sequencing. Plasmid DNA was then introduced into *S. aureus* by electroporation and the cells were plated on TSA plates supplemented with 5µg/ml erythromycin and 200µg/ml X-gal.

Allelic exchange

Allelic exchange was carried out using the methods described by Arnaud (Arnaud *et al.*, 2004) A single blue colony carrying the appropriate pMAD derivative was picked from TSA plates supplemented with erythromycin (5µg/ml) and Xgal (200 µg/ml), transferred to TSB supplemented with erythromycin (5 µg/ml), and incubated overnight with shaking at 44°C to select for cointegration of the plasmid. The overnight culture

was diluted in fresh medium and grown for about 6 hours at 44°C with shaking. Serial dilutions were then prepared and plated on pre-warmed TSA supplemented with erythromycin and Xgal and incubated at 44°C for 48 hours. Several isolated blue colonies were then inoculated into TSB and incubated with shaking at 30°C overnight to resolve the cointegrate. The overnight culture was then diluted with fresh medium and allowed to grow at 30°C for about 6 hours. This subculture was added to TSB and incubated overnight at 44ºC to cure cells of the plasmid. The overnight culture was again diluted in TSB and allowed to grow for an additional 6 hours at 44°C. Serial dilutions were prepared and aliquots were spread on TSA plates supplemented with Xgal and incubated for 48 hours at 30°C. Several white colonies were streaked onto TSA plates supplemented with Xgal, and onto TSA supplemented with Xgal and erythromycin to screen for white colony phenotype and sensitivity to erythromycin, indicating loss of the plasmid. Colony PCR with Taq polymerase and appropriate flanking oligonucleotides was performed to screen for mutants that had the appropriate deletion or insertion. Mutants were further examined for unwanted additional mutations by sequencing the amplicon of the entire region carried on the allelic exchange plasmid.

Southern blotting

Probe Construction

Digoxigenin-11-2'-deoxy-uridine-5'-triphosphate (DIG-11-dUTP) (Roche Applied Science, Mannheim, Germany) labeled probes specific for 80α and SaPI1 were prepared for analysis of Southern blots. Dig-11-dUTP was randomly incorporated into the PCR product in place of dTTP. PCR conditions were the same as above for Taq except for the following changes: in place of the standard dNTP mix, a dig-dNTP mix

consisting of 0.2 mM each dATP, dCTP and dGTP, and 0.13 mM dTTP and 70 µM DIG-11-dUTP was used in each PCR.

Preparation of g-DNA

Freshly prepared BHI media was inoculated with a previously grown overnight broth culture of a *S. aureus* strain at a ratio of 1:100 and incubated at 30ºC on an orbital shaker at 200 rpm. Cell growth was monitored using either the Klett reader until it reached an optical density (OD) of 0.6 (Klett = 50). Mitomycin C was added to terminate lysogeny. The culture was monitored for 1 hour and spun down to collect the cell pellet. The cells were resuspended in 1µl TES [10 mM Tris-HCl (pH 7.5),1 mM EDTA (pH 8.0), SDS(0.1% w/v)], RNaseA (20 µg/ml) and lysostaphin 0. 5 mg/mL. The resuspended cells were allowed to lyse in a waterbath for 1 hour at 37ºC. 20 µl 20% SDS was added and the mixture was further incubated for 30 minutes. 12 µL (3M) NaOAc and 200 µL phenol:chloroform:isoamyl alcohol (25:24:1) were added and the mixture was centrifuged at 14K for 5 minutes. The aqueous phase was collected in a fresh tube and 2 volumes of 95% EtOH were added. The mixture was stored at -20ºC for at least 1 hour. After a period of 1 hour, the mixture was centrifuged at 14k rpm for 5 min to pellet the DNA. Ethanol was carefully decanted from the tube. The DNA was rinsed again with 70% ethanol (500µl) and spun again at 14k rpm for 5 minutes. The ethanol was carefully removed, and the DNA was air dried. The DNA was finally suspended in 20µl of water and resolved on 0.7% agarose gel.

Southern blot transfer

The gels were trimmed and photographed under UV light after staining with ethidium bromide. The resolved DNA was depurinated by incubating the gel in 0.25 N

HCl for 10 minutes, and then rinsed with distilled water. The gel was then soaked in a denaturation solution (1.5 M NaCl, 0.5 N NaOH) for 45 minutes. The gel was removed from the denaturation solution, rinsed with distilled water and soaked in a neutralization solution (1 M Tris-HCl pH 7.4, 1.5M NaCl) for 30 minutes. The gel was again rinsed in distilled water and soaked in fresh neutralization solution for 15 minutes. A positively charged nylon membrane (ICN, Irvine, CA or Roche Applied Science, Mannheim, Germany) and two pieces of Whatman 3MM paper were cut slightly larger than the gel. The nylon membrane was submerged in distilled water until it was completely wet and then transferred to 10XSSC buffer and allowed to equilibrate for five minutes. A reservoir tank filled with 10XSSC was assembled using a glass plate that was wrapped in Whatman 3MM paper. This paper served as a wick during the transfer. The gel was removed from the neutralization solution and placed on center of the wet Whatman 3MM paper. The gel was then carefully surrounded with a parafilm barrier. The nylon membrane was then placed on top of the gel, followed by two pieces of Whatman 3MM paper wetted with 2XSSC. A stack of paper towels about 5 cm thick was placed on top of the Whatman 3MM paper followed by a glass plate and a weight of approximately 500 g. Transfer by capillary action was allowed to proceed for 24 hours, after which the gel and nylon membrane were recovered. Agarose was removed from the nylon membrane by soaking the membrane for 15 minutes at room temperature in 6X SSC. The nylon membrane was then placed on Whatman 3MM paper saturated in 10X SSC and DNA was fixed to the membrane by UV-irradiation in a UV Stratalinker 1200 (Stratagene, La Jolla, CA) at 120,000 µjoules/cm^2. The gel was visualized under UV light to verify that DNA was transferred.

Hybridization and detection

Reagents were prepared using the DIG Wash and Block Buffer Set (Roche Applied Science, Mannheim, Germany) that contains concentrated buffers used in hybridization and detection. The appropriate volume (20 ml/100 cm^2 nylon membrane) of hybridization buffer (5X SSC, 0.1%N-lauroyl sarcosine, 0.2% SDS and 1% blocking solution) was pre-warmed to the hybridization temperature (45°C) and placed into a hybridization bottle. The nylon membrane was placed into the bottle and incubated at 40°C for 30 minutes with gentle agitation. The Dig-11-dUTP labeled probe (3 µl) was placed into a microcentrifuge tube with 50 µl of deionized water and boiled for 5 minutes, then rapidly cooled on ice. The denatured probe was mixed with pre-warmed hybridization buffer (3.5 ml/100 cm^2 nylon membrane). The pre-hybridization solution was decanted and replaced with the hybridization-probe solution and incubated for 16 hours at 45°C. After hybridization, the nylon membrane was washed two times for 5 minutes at room temperature in a high salt buffer (2X SSC, 0.1% SDS) followed by washing two times for 15 minutes each at 65°C in a lower salt buffer (0.1X SSC, 0.1% SDS). After washes, the membrane was placed into a clean tray with 1X washing buffer (0.1 M maleic acid, 0.15 M NaCl pH 7.5, 0.3% (v/v) Tween20) for 5 minutes at room temperature with gentle agitation on an orbital platform. The solution was aspirated from the tray and replaced with 100 ml of 1X Blocking solution for 30 minutes with gentle agitation. Subsequently, the Blocking solution was replaced with (20 ml) of 1X Blocking solution containing Anti-digoxigenin-AP (75 mU/ml) and incubated for 30 minutes with gentle agitation. The antibody-containing solution was removed and the membrane was equilibrated with 20 ml of 1X Detection buffer for 5 minutes with gentle agitation. The

membrane was placed into an opened hybridization bag and 1 ml of CSPD working solution (diluted 1:100 with 1X Detection buffer) was applied to the membrane. The solution was spread evenly across the membrane by closing the hybridization bag and incubated at room temperature for 5 minutes. The excess liquid was squeezed from the bag, which was then sealed and incubated for 15 minutes at 37°C. The membrane was exposed to X-ray film (Kodak, Rochester, NY) and developed in an X-O-Mat (Kodak, Rochester, NY).

Expression of proteins in *E. coli*

Expression of 80α major capsid protein and scaffolding protein from pPD2 in *E. coli* strain BLR(DE3) (Invitrogen) was accomplished by induction with 0.5 mM IPTG at 37 °C. The cells were harvested 2 h post induction, lysed by passage through an EmulsiFlex C3 high-pressure homogenizer (Avestin, Inc., Ottawa) and prepared for SDS-PAGE.

Expression of proteins in *S. aureus*

Overnight cultures of respective *S. aureus* strains containing pG164 derivatives were inoculated in freshly prepared Antibiotic Medium 3. The cells were allowed to grow till mid exponential phase (Klett 60). 0.5mM IPTG was added and cells were harvested 2 hours post induction. Cells were pelleted by centrifuging at 4000 rpm for 10 minutes and resuspended in PBS buffer. Cells were then either mechanically disrupted by using a mini bead beater or boiled in XT reducing buffer (Biorad Hercules, CA) for 10 minutes and subjected to SDS-PAGE analysis

Polyacrylamide Gel Electrophoresis

Protein samples were analyzed using the Criterion XT protein system (Biorad, Hercules, CA) and were separated on either pre-cast Bis-Tris 12% gels or 4-12% gradient gels (Biorad Hercules, CA). 4X XT loading buffer was added to samples to achieve a 1X concentration and these samples were then heated for 10 minutes prior to gel loading. The proteins were separated under constant voltage at 200V for 50 minutes in 1X MOPS (Bio-Rad, Hercules, CA) running buffer. A Precision Plus Protein Standard dual color marker (10 kDa-250 kDa) (Bio-Rad, Hercules, CA) was run with samples to estimate the molecular weight of the protein. After electrophoresis, the gels were rinsed in deionized water three times for 5 minutes each and then stained with Bio-Safe™ Coomassie (Bio-Rad, Hercules, CA) followed by destaining in water overnight. The bands were visualized under visible light and images were taken using Alpha imager system.

Table 2. Oligonucleotides used in this study

Primer	Sequence (5'–3') restriction sites indicated in bold
	pPD8
PKD 9	GA**G GAT CC**G GTC GAA AAC AAG GAC TTT AGC GAT AGA G
PKD 12	GA**G GAT CC**C AAT GAT TTC GGG CAT GTT ACC ACT CC
PKD10	ATG CCT CCG TTA ATT TTT AAT AAT TCT ATT TTC TTC CAT GAG ATA TAC CTC CAT TTA TAG TCT GTC
PKD 11	ATG GAA GAA AAT AGA ATT ATT AAA AAT TAA CGG AGG CAT TTA AAT GGA ACA AAC
	pPD11
SMT51	ACA**CCATGG**GCATACAGATATTCTCTGGA
SMT52	ACA**AAGCTT**GTGGATGATATACCGTTAGAG
SMT53	ACA**AAGCTT**CGCTTGTTTTGCCGTTAA
SMT54	ACA**CCATGG**CAATATGCAGGAGATTTCAAG
	pPD14
SMT110	ACA**GGATCC**GGAAGTATATTGTCGGGCAAC
SMT113	ACA**GGATCC**ATGAGTATGAGCCACTCGC
PKD19	TCC TTA ATA TTC GAC GGT TAC CAT GCG TCT CGC
PKD20	CGC ATG GTA ACC GTC GAA TAT TAA GGA GTG TTA AAA ATG CCG
	pPD17
PKD45	AA**G GAT CC**A ACA CAG TTG AAC GTA ATA GTA TCG CA
PKD48	AA**G GAT CC**T GTT TGT TCC ATT TAA ATG CCT CCG T
PKD46	ATC AAT CAC CGC TTT TCA CAT CTC GCT TTA ACT CAT TCT CGA TTG CTT
PKD47	AGA ATG AGT TAA AGC GAG ATG TGA AAA GCG GTG ATT GAT TTA AAA GTA AAG TTT
	pPD34
PKD93	TAC ATA CTT TCA ATT TG**G AAT TC**A CAC ATC
PKD94	CGC **GGA TCC** AGT ATC GCA GAA AAA GAG ACG TAT C
PKD42	AAA ATG AAA ACT GAA TCG TAC CGT TAA AAA GAA AGG GTA ATT AAA TGG AAA CAA AAT
PKD43	AAT TAC CCT TTC TTT TTA ACG GTA CGA TTC AGT TTT CAT TTT TAT CCC CTT
	pPD35
PKD91	GAC GC**G GAT CC**T CGT GCA ATC TTA CTG TTT TCA ATT G
PKD92	CAT G**CC ATG G**AC TGA ATC GTA CTT TAA AGA ATA CAA
PKD38	GGT AAT TAA ATG GAA ACA AAA AGG ATT ATT AAA TAG CAA TGA TTG CCT ATC CAA
PKD39	TCA TTG CTA TTT AAT AAT CCT TTT TGT TTC CAT TTA ATT ACC CTT TCT TTT TAA CGT
	pPD27
PKD64	AA**G GAT CC**A TGG AGG CAG AAC CTT TTA TG
PKD65	AA**G GAT CC**T ACT CTT TCG CAT ACA TTC GAT AGT TAT A
PKD54	TAT ACC TCT CCC ATA TTC GCC ATT CAA TAT TCC TCC
PKD55	ATG GCG AAT ATG GGA GAG GTA TAA AAA TAG GGC GTT A
	pPD36
PKD-72	GGT T**GG ATC C**TT ATG CAA CAG CTC AAA CTG AAT TAT CTA
PKD 97	TTA ATT ACC CTT TCT TTA ACT AAC TTG CCT TTA AAA AAC TTT ACT TTT AAA TC
PKD 98	TTA AAG GCA AGT TAG TTA AAG AAA GGG TAA TTA AAT GGA AAC AAA ATA CG
PKD-75	ATT TAA TTT ACT GTC ATA CAA CTA TTT AAT AAT CCT GTT TTG CTT AGC TAA ATT TTG
PKD-76	CAA AAC AGG ATT ATT AAA TAG TTG TAT GAC AGT AAA TTA AAT GTT TGG AGG ATA CTA
PKD-77	GGT TT**G GAT CC**C TGA ACG TAA TTG TTC CAG CTC T
	pPD37
PKD-72	GGT T**GG ATC C**TT ATG CAA CAG CTC AAA CTG AAT TAT CTA
PKD-73	ATC CCC TTT ACA CTT AAC TAA CTT GCC TTT AAA AAA CTT TAC TTT TAA ATC

PKD-74	TTT TTT AAA GGC AAG TTA GTT AAG TGT AAA GGG GAT AAA AAT GAA AAC TG
PKD 99	ATT TAA TTT ACT GTC ATA CAA TTA ACG TTT TAA AAA CAA CTT GTT ATT GTG TTC GTA
PKD 100	AAG TTG TTT TTA AAA CGT TAA TTG TAT GAC AGT AAA TTA AAT GTT TGG AGG ATA CTA
PKD-77	GGT TT**G GAT CC**C TGA ACG TAA TTG TTC CAG CTC T
pPD28	
PKD-72	GGT T**GG ATC C**TT ATG CAA CAG CTC AAA CTG AAT TAT CTA
PKD-73	ATC CCC TTT ACA CTT AAC TAA CTT GCC TTT AAA AAA CTT TAC TTT TAA ATC
PKD-74	TTT TTT AAA GGC AAG TTA GTT AAG TGT AAA GGG GAT AAA AAT GAA AAC TG
PKD-75	ATT TAA TTT ACT GTC ATA CAA CTA TTT AAT AAT CCT GTT TTG CTT AGC TAA ATT TTG
PKD-76	CAA AAC AGG ATT ATT AAA TAG TTG TAT GAC AGT AAA TTA AAT GTT TGG AGG ATA CTA
PKD-77	GGT TT**G GAT CC**C TGA ACG TAA TTG TTC CAG CTC T
pPD1	
SMT34	GACT**CATATG**GAACAAACACAAAAATTAAAAT
SMT35	GACT**GGATCC**TTATTAAACTTCTCCTGGAACT
pPD2	
SMT61	TCAG**GGATCC**GAAGGAGATATCTCATGGAAGAAAATAAACTTAAG
SMT62	CTTC**GTCGAC**TTAAATGCCTCCGTTAATTTTTAA
pPD18	
PKD57	AA**C TGC AG**A ACG GAG GCA TTT AAA TGG AAC
PKD58	AA**C TCG AG**A AGT CAG GCG CGC CAA TTG TTT ATT AAA CTT CTC CTG GAA CT
pPD20	
PKD 78	CAC A**GG ATC C**AT GGA ACA AAC ACA AAA ATT AAA ATT AAA TTT GC
PKD 58	AA**C TCG AG**A AGT CAG GCG CGC CAA TTG TTT ATT AAA CTT CTC CTG GAA CT
pPD21	
PKD 79	CAC A**GG ATC C**AT GGA AGA AAA TAA ACT TAA GTT TAA TTT GCA A
PKD 58	AA**C TCG AG**A AGT CAG GCG CGC CAA TTG TTT ATT AAA CTT CTC CTG GAA CT
pPD22	
PKD 66	TTT T**GG ATC C**AT GTT AAA AGT AAA CGA ATT TGA AAC AGA TAC A
PKD 67	AAA A**CT GCA G**TA ATT GTT TAT TAA ACT TCT CCT GGA ACT G
pPD44	
PKD 86	GGC AAA **GGC GCG CC**A AGA AAG GGT AAT TAA ATG GAA ACA AAA TAC GA
PKD 87	GGC AA**C TCG AG**T TGC TAT TTA ATA ATC CTG TTT TGC TTA GCT AAA TTT
pPD45	
PKD-88	GGC AAA **GGC GCG CC**A AGG GGA TAA AAA TGA AAA CTG AAT CGT AC
PKD 87	GGC AA**C TCG AG**T TAA CGT TTT AAA AAC AAC TTG TTA TTG TGT TCG
pPD46	
PKD 88	GGC AAA **GGC GCG CC**A AGG GGA TAA AAA TGA AAA CTG AAT CGT AC
PKD 87	GGC AA**C TCG AG**T TGC TAT TTA ATA ATC CTG TTT TGC TTA GCT AAA TTT
pPD47	
PKD-88	GGC AAA **GGC GCG CC**A AGG GGA TAA AAA TGA AAA CTG AAT CGT AC
PKD 85	GGC AA**C TCG AG**T TAC ACC ACT TTT ACA TAT GAA GAT TGG TG
pPD51	
PKD-88	GGC AAA **GGC GCG CC**A AGG GGA TAA AAA TGA AAA CTG AAT CGT AC
PKD 87	GGC AA**C TCG AG**T TAA CGT TTT AAA AAC AAC TTG TTA TTG TGT TCG
pPD52	
PKD 88	GGC AAA **GGC GCG CC**A AGG GGA TAA AAA TGA AAA CTG AAT CGT AC
PKD 85	GGC AA**C TCG AG**T TAC ACC ACT TTT ACA TAT GAA GAT TGG TG

Chapter 3

The structural characteristics of 80α and SaPI1 capsids and procapsids

Introduction

All phages must, in order to propagate, exit their host and face the harsh outside environment. They need to protect their nucleic acid and also have a mechanism to infect their next host. The nucleic acid is packaged in a proteinaceous coat that is stable enough to withstand the rigors of the outside environment, which include pH changes, extremes of temperature, salinity, pressure and radiation levels. This protein coat is a macromolecular assembly of smaller subunits known as capsid proteins encoded by a relatively short nucleic acid sequence and can encapsidate a relatively large volume of nucleic acid. This strategy of using assembly of multimeric capsid proteins circumvents the problem of encoding large capsid subunits and thereby helps the phage maintain genetic economy. Of the 5 known platonic solids in nature, the icosahedron generates the maximum enclosed volume for shells comprised of a given size subunit (Caspar *et al.*, 1962) and hence it is not surprising that most phages have shells that conform to icosahedral symmetry. The discussion in this chapter is limited only to structural characteristics of icosahedra since both 80α and SaPI particles have isometric icosahedral symmetries.

The icosahedron

An icosahedron is an isometric structure with 12 pentagonal vertices and 20 identical equilateral triangular faces related by two-, three-, and fivefold axes of rotational symmetry (Fig 6.a). Any regular icosahedron has six five-fold axes that pass through the twelve vertices, ten threefold axes that pass through the centroids of twenty equilateral triangular faces and fifteen twofold axes through the edges. Thus this defined set of symmetry elements, belongs to the icosahedral symmetry point group (I_h) (5 3 2) and generates 60 identical repeat units. This point group can be used to describe any icosahedral structure. Application of these symmetry elements to any subunit that does not lie on a symmetry axis causes it to be repeated 60 times in the complete structure. The asymmetric unit (Fig 6.b) of an icosahedron is defined as the wedge-shaped volume that extends from the icosahedron's center along edges formed by a threefold axis and two adjacent fivefold axes (Baker *et al.*, 1999). The simplest icosahedral structure is the one that has 60 identical subunits that exist in a structurally identical environment. However, construction of a capsid based on such a simple icosahedral design severely restricts the interior volume available for packaging the nucleic acid genome. In fact, most icosahedral capsids are composed of more than 60 subunits and these subunits cannot have identical environments. Caspar and Klug in 1962 put forth two important concepts, namely the theory of quasi–equivalence and the concept of triangulation numbers (Caspar *et al.*, 1962). They suggested that the "shell is held together by the same types of bonds throughout, but that these bonds may be deformed in slightly different ways in the different, non-symmetry related environments. Molecular structures are not built to conform to exact mathematical concepts but rather

to satisfy the condition that the system be in a minimum energy configuration." In other words, shells comprised of more than 60 subunits would be formed from chemically equivalent subunits, but with slight changes in their bonding.

Triangulation numbers

The second important concept introduced by Caspar and Klug to explain structures of icosahedral shells was that of triangulation numbers (Caspar *et al.*, 1962). Triangulation number is a geometric and abstract concept that refers to the organization of icosahedral structures and does not necessarily strictly correspond to the protein subunits of the virus. An isometric shell can be built by starting with a flat array of hexagons which is defined by a hexagonal coordinate system with two oblique axes h and k (Goldberg,1937; Luczak *et al.*, 1976) (Fig 7a). All internal bonds are identical in environment in this flat hexagonal net. In order to generate a regular closed structure out of this flat net, one needs to curve it at specific places. This can be done by converting some of the hexagons to pentagons in this array. Changing hexagons to pentagons not only introduces curvature but also introduces the concept of quasi-equivalence. These pentagons are interchanged with specific hexagons governed by the triangulation number (T) and related to the equation: $T = h^2 + k^2 + hk$ where h, k are the coordinates of the next pentagon on the hexagonal coordinate system. Thus for obtaining an icosahedral structure with a triangulation number T=4, the hexagon with coordinates (2,0) at its center would be converted to a pentagon (Fig 7b). Similarly, for T=7, converting either of the hexagons with coordinates (2,1) or (1,2) to pentagons would yield icosahedral structures with T=7l (Fig 7c) or T=7d, respectively. In other

Figure 6. Axes of symmetry and asymmetric unit of an icosahedral structure

(A) A schematic representation of an icosahedral structure with two fold, three fold and fivefold axes of symmetry. The three fold axis is shown by a triangle; the twofold axis with an oval and the five- fold axis with a pentagon.
(B) An icosahedral structure showing an asymmetric unit defined by a white triangle which is bounded by the lines joining adjacent fivefold and three fold positions. The numbers, 2,3 and 5, indicate the corresponding axes of symmetry.

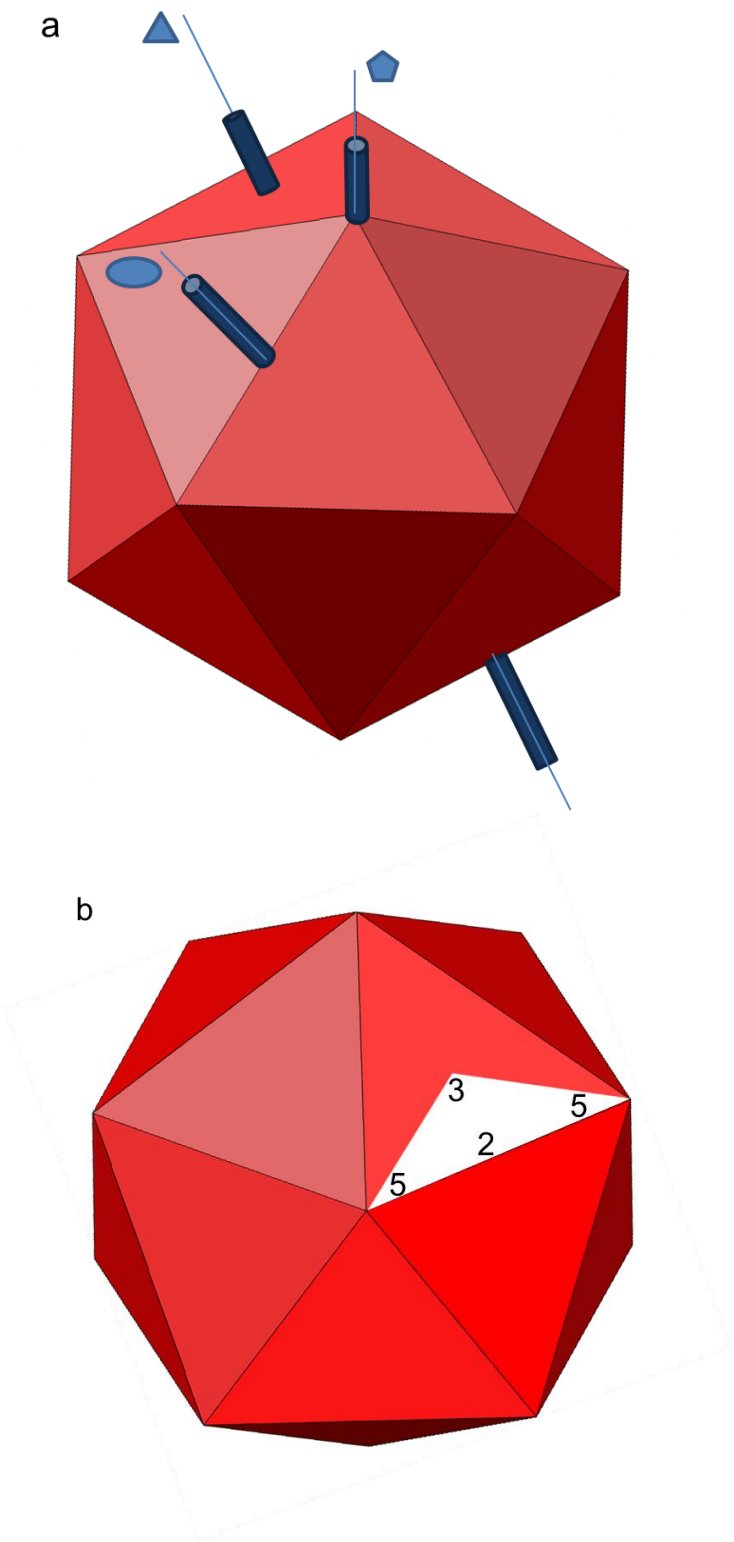
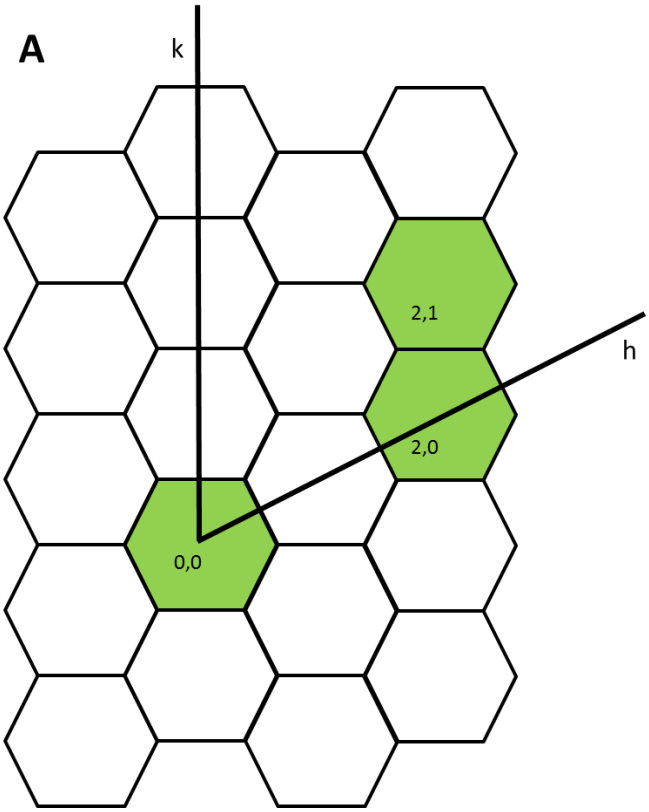
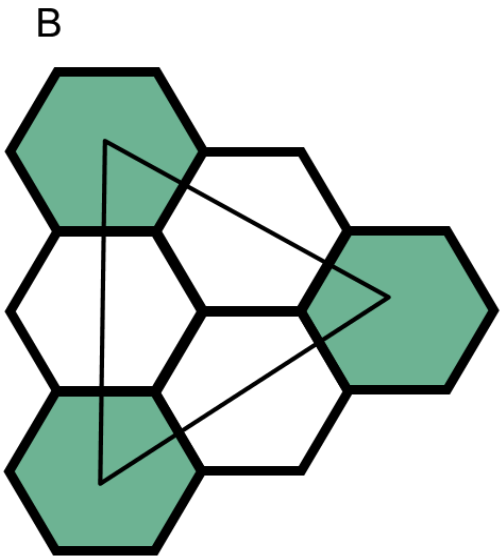
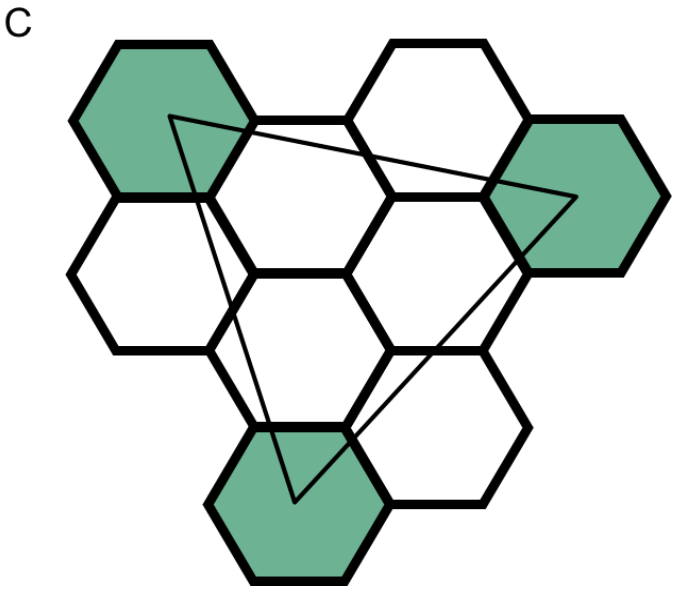

a
b
3
5
2
5

Figure 7. Geometric principles of constructing icosahedral lattices of defined triangulation number.

(A) hexagonal coordinate system showing an origin and two oblique axes defined by h and k. (B and C) Relevant hexagons that would be converted to pentagons for T=4 and T=7 are shown in green.


A
k
h
2,1
2,0
0,0


B


C

words, the triangulation number denotes the surface area of one triangular face of the icosahedral structure. The number of quasiequivalent environments increases as the triangulation number increases. For example, a T=4 capsid has four different environments and a capsid with T=7 has seven. Also, though not strictly true, the triangulation number indicates the number of subunits present in the capsid and is given by the relation: number of subunits, N= 60(T).

Rationale

In order to understand why smaller sized capsids are predominantly formed in the presence of SaPI1, it was necessary to investigate the structural characteristics of both types of particles and study their similarities and differences. Insight into their shape, size and icosahedral symmetry would yield valuable information about the principles governing the self-assembly process, maturation and the structure of capsids. Three dimensional reconstructions of icosahedral viruses from cryo-electron micrographs are now being increasingly used to complement X-ray crystallographic studies (Tang *et al.*, 2002; Rossmann *et al.*, 2005) and sometimes even to supplement them (Topf *et al.*, 2008) to yield valuable information about the triangulation numbers, the number of subunits in the capsid shell and information concerning the alteration of subunit conformations and their interactions in the structure. Extensive studies at nanometer resolution levels have been carried out in HK97 (Conway *et al.*, 1995; Lata *et al.*, 2000; Gertsman *et al.*, 2009), P22 (Nemecek *et al.*, 2007; Lander *et al.*, 2009; Teschke *et al.*, 2010; Chen *et al.*, 2011; Tang *et al.*, 2011) and T4 (Kanamaru *et al.*, 2004; Kostyuchenko *et al.*, 2005; Leiman *et al.*, 2010; Luque *et al.*, 2010; Rao *et al.*, 2010) that have generated a wealth of information regarding the mechanisms involved

in self-assembly processes, protein-protein interactions and conformational changes in the tertiary structures of the capsid protein during the maturation process. We wanted to carry out similar studies with 80α and SaPI1 capsids and procapsids in order to understand the mechanisms involved in capsid size redirection.

Generation of 80α and SaPI1 procapsids and capsids

In *Caudovirales*, the first stage of the capsid self-assembly pathway is the formation of a capsid precursor. This capsid precursor, also known as a procapsid, is formed by co-polymerization of multiple copies of major capsid and scaffold proteins with a dodecameric portal complex at one of its vertices. Multiple copies of scaffold proteins direct the polymerization of major capsid protein to form a properly dimensioned shell. The portal complex, usually a dodecamer of portal protein subunits, occupies one of the five-fold vertices of the procapsid and ultimately serves as a route for entry and exit of DNA into the capsid and also as a connector between the head and tail. The procapsid is a transient intermediate and undergoes rapid maturation through a series of steps. For example, in a study done on phage HK97 the half-life of procapsids was found to be approximately 30 seconds (Lata *et al.*, 2000). The process of maturation involves scaffolding protein removal, usually through its proteolytic cleavage. Proteins that are only transiently involved in the self-assembly process would be completely missed during our analysis on mature particles. Therefore we wanted to analyze and compare the protein composition of procapsids of both sizes by SDS PAGE and mass spectroscopy. Since procapsids are transient intermediates and their maturation is highly efficient, isolation of procapsids at significant levels posed a technical problem. Arrest of the self-assembly process at the procapsid stage, in theory,

could be achieved by disrupting one of the maturation steps. Since procapsid maturation is tightly linked to DNA packaging (Earnshaw *et al.*, 1980), the procapsid stage could also be arrested by mutations affecting DNA packaging. Thus in theory, mutations in large terminase or small terminase genes would affect DNA packaging and arrest the self-assembly process at the procapsid stage. We created mutants defective in DNA packing by constructing in-frame deletions of the phage and SaPI1 small terminase subunits, namely ST24 and ST37 (Poliakov *et al.*, 2008). ST24 is a strain containing an 80α prophage with its *terS* deleted while ST37 contains a wt 80α prophage and SaPI1 with its *terS* deleted. We found that induction of strain ST37 did not yield enough SaPI1 procapsids for detection, due to residual nonspecific packaging caused by 80α TerS, and hence we created a new strain, ST63, with both the phage and SaPI1 small terminase subunits deleted. Procapsids were purified by equilibrium sedimentation on CsCl gradients, as described in Materials and Methods, and then further purified by velocity sedimentation in sucrose gradients and used for analysis by electron microscopy and mass spectrometry.

Results

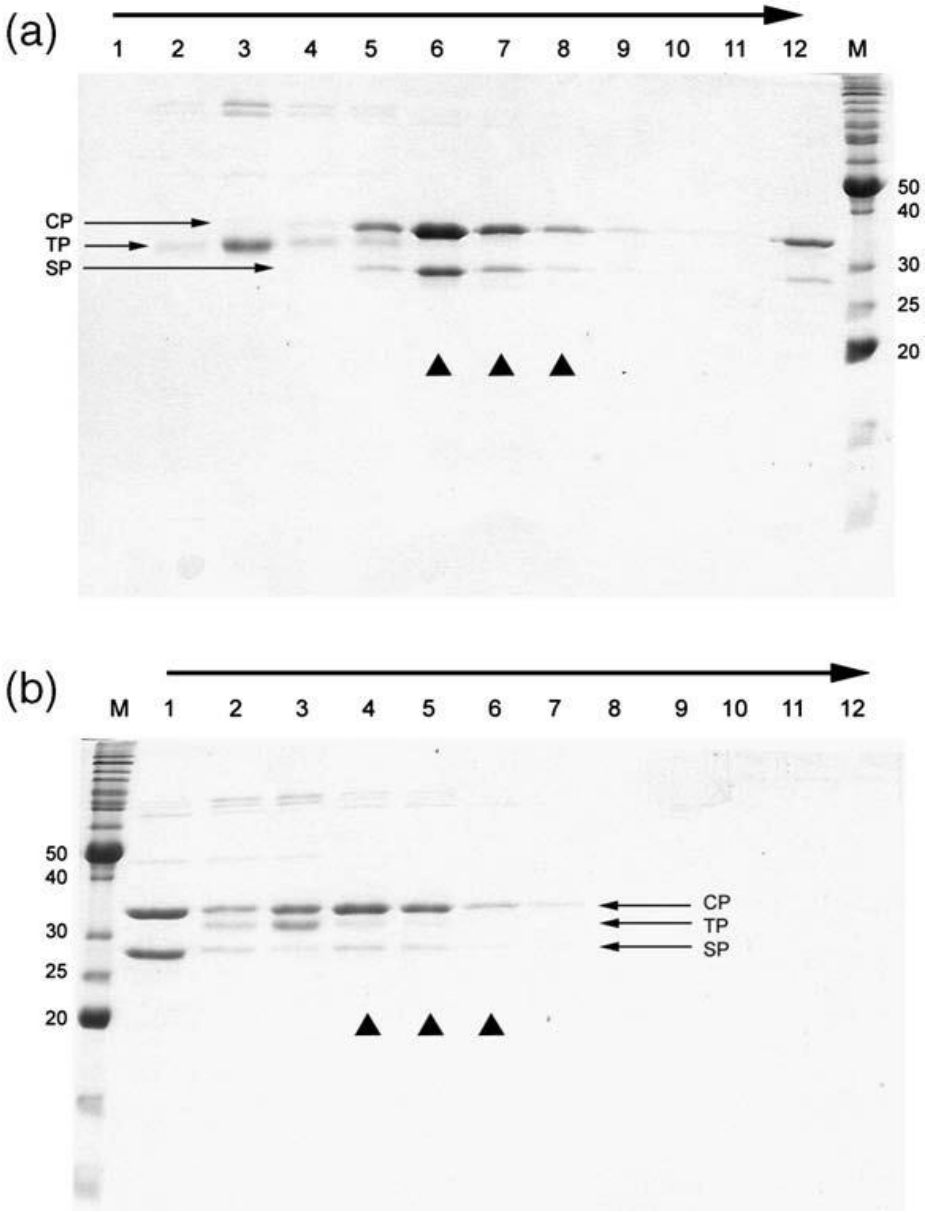
SDS-PAGE analysis of sucrose gradient fractions

80α and SaPI1 procapsid production was induced in strains ST24 and ST37 respectively, and procapsids were purified on CsCl gradients followed by sucrose gradients as described previously. Unlike mature virions, procapsids cannot be separated from tails and other proteins only by equilibrium sedimentation. SDS-PAGE analysis of procapsids generated from the lysogenic 80α *ΔterS* mutant strain ST24 and

Figure 8. SDS-PAGE of sucrose-gradient-separated procapsids

80α (a) and SaPI1 (b) procapsids. Fraction numbers (1-ml fractions) and the direction of sedimentation (arrow) are indicated. M, marker; molecular mass, as indicated (kDa). The positions of the capsid protein (CP), major tail protein (TP) and scaffolding protein (SP) are indicated by arrows, while the triangles indicate fractions that were pooled and used for cryo-EM and MS analyses.

Reprinted from J Mol Biol. 380 (3) Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids, 465-475, (2008) with permission from Elsevier


(a)
1 2 3 4 5 6 7 8 9 10 11 12 M
CP
TP
SP
50
40
30
25
20
(b)
M 1 2 3 4 5 6 7 8 9 10 11 12
50
40
30
25
20
CP
TP
SP

fractionated on sucrose gradients showed that tail proteins sedimented more slowly and were found in fractions F3 and F4 (Fig. 8a), while the major capsid and scaffolding proteins were found mainly in several faster sedimenting fractions (F5–F8). Comparison of band densities of scaffold and major capsid in the SDS-PAGE gel showed that the molar ratio of scaffolding to capsid protein (SP:CP) varies between 0.33 (lane 8) and 0.76 (lane 6). A negligible amount of soluble protein was seen at the top of the gradient, suggesting that structural proteins were efficiently incorporated into stable procapsids. The fractions (F6–F8) containing mainly capsid and scaffolding protein were pooled, concentrated by pelleting and observed by cryo-EM (Fig. 9 and 10).

Structural characteristics of 80α and SaPI1 procapsids

EM analysis of 80α and SaPI1 sucrose gradient purified samples showed the presence of particles that were distinctly different than the mature virions. The particles were smaller in size than their respective mature capsids. They also had a more rounded shape, thicker walls and an internal core consisting of radial segments, consistent with features of other bacteriophage procapsids. The formation of procapsids in the *terS* mutants clearly established that packaging is tightly linked to capsid maturation and blocking the packaging function also blocks capsid maturation. The internal core of both procapsids consisted of radial segments approximately 12–14nm long (Fig 9b), which would correspond to 80–100 residues of an α-helix. Mature 80α scaffold has193 amino acids and is predicted to be predominantly α-helical, like other known scaffolding proteins. It could presumably, like P22 scaffold, interact with the major capsid protein through both its termini (Chen *et al.*, 2011).

80α capsids

Cryo-electron microscopy and 3-D reconstruction were done in the laboratory of our collaborator, Dr. Terje Dokland (Poliakov *et al.*, 2008; Spilman *et al.*, 2011). 80α virions have capsids with 57 nm diameter (side to side) and 63 nm diameter (vertex to vertex) and with 190 nm long flexuous tails (Figs 9 and 10). The star-shaped, 6-fold symmetric baseplates can also be seen. The interior of the capsids have a punctuate appearance with a spacing of 2.5nm that is due to the closely packed DNA inside. Symmetric 3-D reconstruction of 80α capsids to a resolution of 1.02 nm shows that 80α has a *T=7l* icosahedral symmetry corresponding to 420 copies of 80α gp47 organized into 60 hexamers that are located on the flat triangular faces of the icosahedra and 12 pentamers located on the icosahedral fivefold vertices (Fig 11a and b). In an asymmetric virion, one of the pentamers is replaced by a portal complex. The icosahedral symmetry of T=7 suggests that gp47 exists in seven quasi-equivalent environments, as shown schematically in Fig 11. The capsomers seem to interact through three types of trivalent contacts that are referred to as type1 (EEE), type 2 (ABG), and type 3 (CDF) (Fig 12). Much of the capsid density is clustered in distinct trefoils delineated by 2-3 nm holes at the contact points. The hexameric connections are less extensive and characterized by a hole with a 1.2 nm diameter in the center. The cross-section of the capsid shows that the shell is, on average, 2.4 nm thick and has a flat appearance except at the fivefold vertices (Fig 11c) and there are 2 nm holes at the 2-fold axes. Furthermore, at least eight concentric layers of density with a spacing of 2.5 nm are visible inside the shell. This density corresponds to the packed DNA.

Figure 9. EM analysis of 80α particles

(a) Cryo-EM of sucrose gradient-purified 80α procapsids. Inset: A 2× magnified view of one procapsid, with dimensions of the shell and the inner core indicated.
(b) Cryo-EM of CsCl-purified 80α virions. The 2-nm spacing of the internal DNA is clearly visible. One thin-walled empty capsid is also shown (inset).
(c) Negatively stained CsCl-purified 80α procapsid fraction containing a mixture of procapsids, tails and procapsids with attached tails.
(d) Cryo-EM of sucrose gradient- purified SaPI1 procapsids. Inset: One of the about 5% 80α-sized procapsids found in the SaPI1 procapsid sample. The scale bar represents 100 nm.
Reprinted from J Mol Biol. 380 (3) Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids, 465-475, (2008) with permission from Elsevier

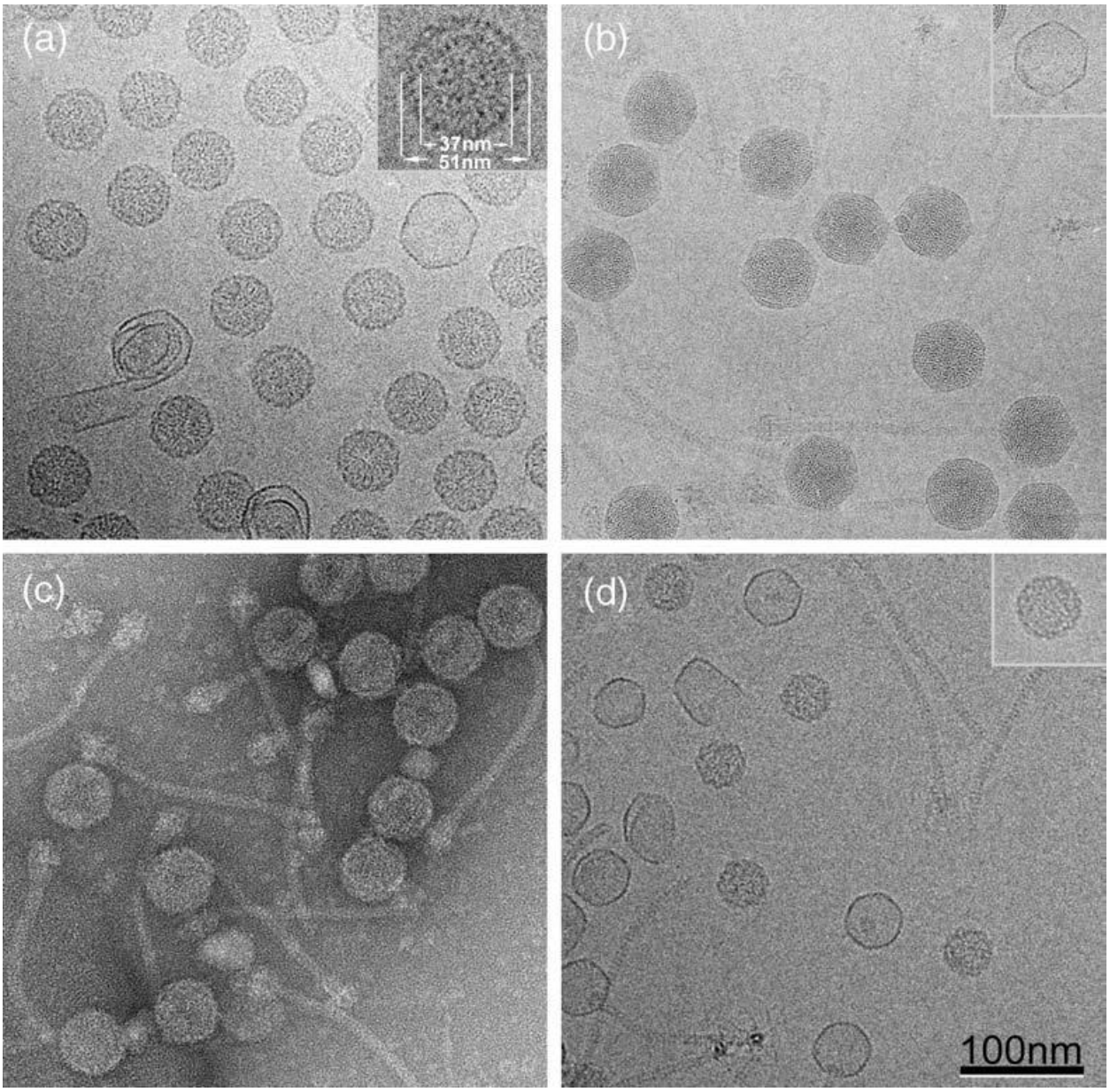

(a)
37nm
51nm
(b)
(c)
(d)
100nm

Figure 10. Cryo-electron micrographs of 80α.

(a) Virions and (b) procapsids. A baseplate seen end-on is indicated (arrow), showing the hexagonal star-like structure with six appendages. The scale bar represents 100 nm.
Reprinted from J Mol Biol. 380 (3) Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids, 465-475, (2008) with permission from Elsevier

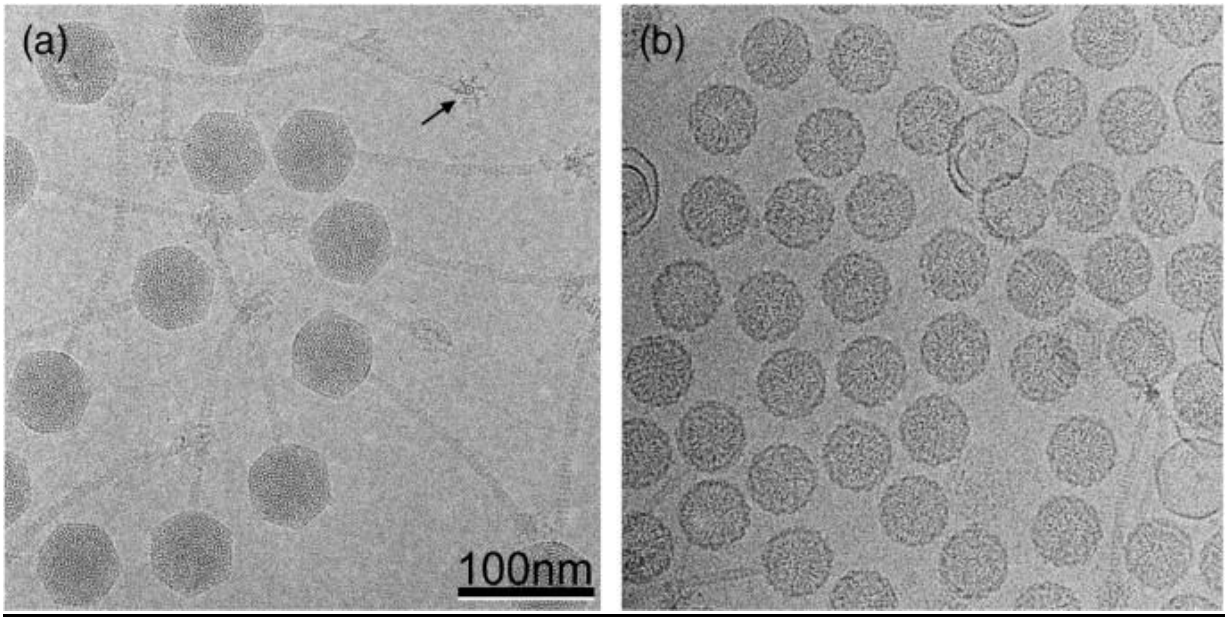

(a)
100nm
(b)

Figure 11. Icosahedral reconstructions of 80α capsids.

(a) Schematic diagram showing one triangular face (large triangle) delimited by three 5-fold and three 2-fold axes (pentagons and ovals, respectively). The icosahedral threefold (filled triangle) is in the center of the face. Each gp47 subunit is shown as a sphinxlike shape, representing the A domain, P domain, E loop, and P loop as shown in the inset. The seven gp47 subunits corresponding to one asymmetric unit are colored: A, red; B, blue; C, teal; D, cyan; E, green; F, yellow; G, orange. The triangular face has an A_5 pentamer at each corner and includes three neighboring BCDEFG hexamers. One neighboring hexamer on an adjacent face is also drawn (gray lettering) to show the three different types of trivalent subunit interactions (numbered triangles). (b) Isosurface representation of the 80α virion reconstruction, viewed down a 3-fold axis and rendered at a cutoff level consistent with the calculated mass of the capsid. The capsid is radially colored from red (center of capsid) to blue (farthest from center). One icosahedral face corresponding to the triangle in (a) is shown, with 2-fold, 3-fold, and 5-fold symmetry axes indicated. The A, B, C, D, E, F, and G subunits in one asymmetric unit are marked. (c) Central section through the virion reconstruction in grayscale representation from white (lowest density) to black (highest density). The arrows indicate the 2-fold axes. Pertinent dimensions are indicated. (d) Isosurface representation of the 80α procapsid reconstruction, viewed, rendered, and labeled as in (b). (e) Central section through the procapsid reconstruction, viewed and labeled as in (c).

Reprinted from J Mol Biol. 405 (3) Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids., 863-876, with permission from Elsevier

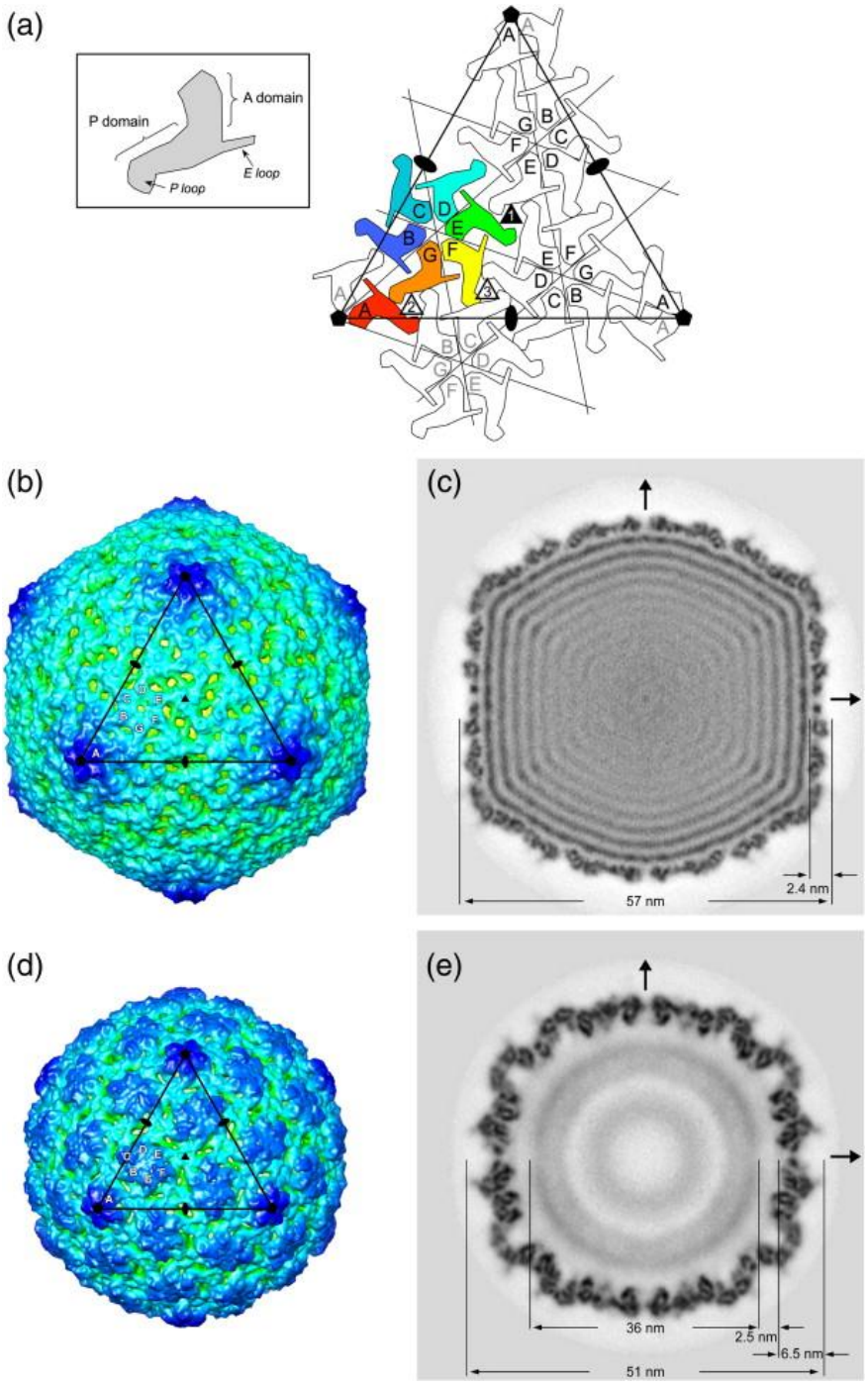
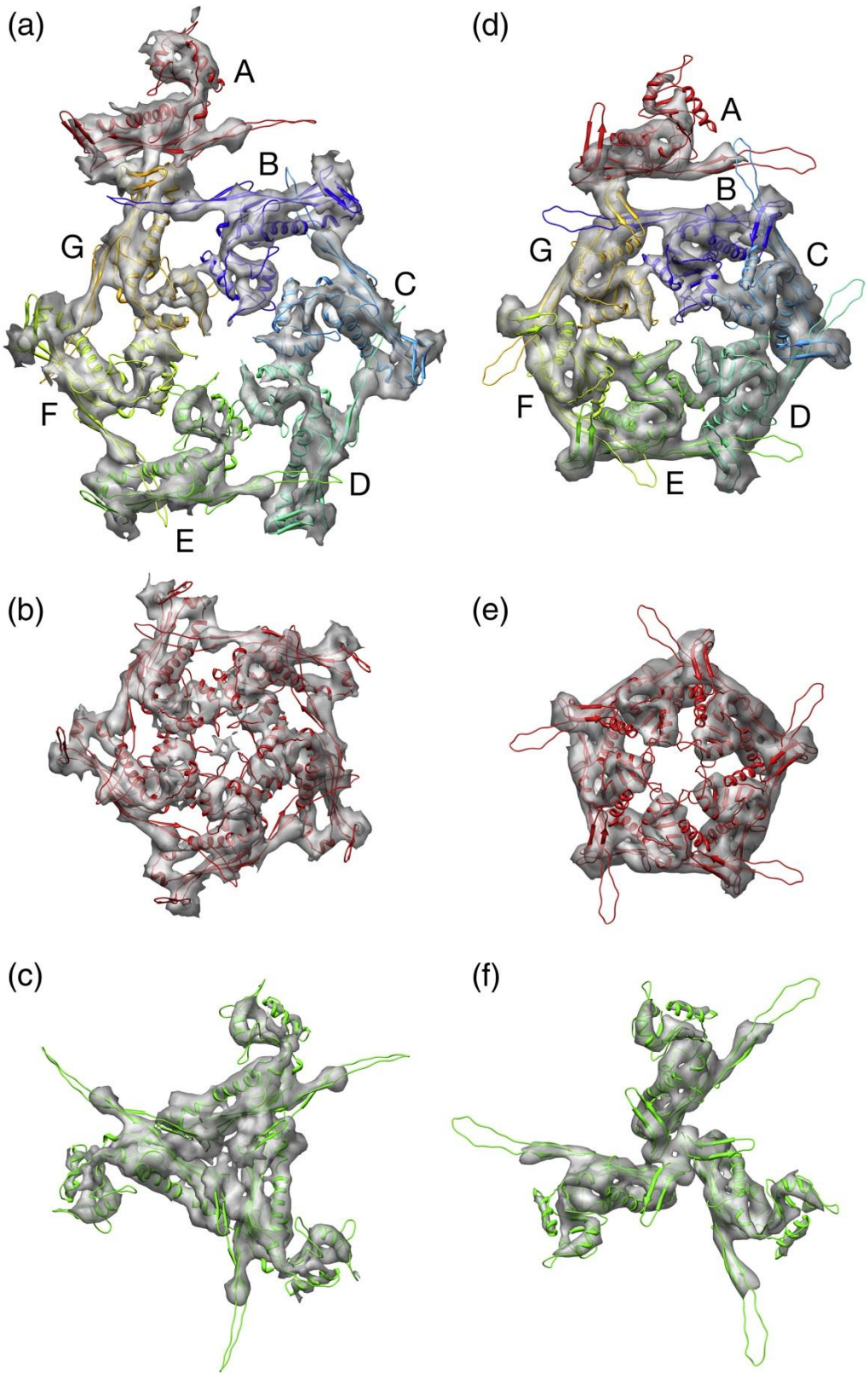

(a)
P domain
A domain
E loop
P loop
(b)
(c)
57 nm
2.4 nm
(d)
(e)
36 nm
2.5 nm
6.5 nm
51 nm

Figure 12. Capsid protein interactions.

Ribbon diagrams of the gp47 model are shown fitted into the capsid electron density in the virion [(a)–(c)] and in the procapsid [(d)–(f)]. (a and d) Asymmetric unit, comprising one-fifth of a pentamer and a BCDEFG hexamer; (b and e) A_5 pentamer and (c and f) the type 1 (EEE) trimer. Subunits are colored as in Fig. 11a.

Reprinted from J Mol Biol. 405 (3) Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids., 863-876, with permission from Elsevier


(a)
A
B
G
C
F
D
E
(d)
A
B
G
C
F
D
E
(b)
(e)
(c)
(f)

80α procapsids

80α procapsids are rounded particles 51 nm in diameter. The procapsid shells are about 4nm in thickness and have an inner core 37 nm in diameter that consists of punctuate density corresponding to scaffolding protein, gp46. The inner core is separated from the shell by a gap of approximately 2nm. Only a small number of spontaneously expanded shells were found, suggesting that maturation of capsids is tightly linked to DNA packaging in the procapsid. Symmetric reconstruction of the 80α procapsid to a resolution of 0.88 nm reveals a roughly spherical shell with an icosahedral symmetry of T=7*l*. The pentamers and hexamers in the procapsid protrude more than those of the mature capsid (Fig 11d). The hexamers have an elongated and skewed appearance as is found in procapsids of other phages (Dokland *et al.*, 1993; Conway *et al.*, 1995; Parent *et al.*, 2010). The protein subunits are closely packed together in the procapsids compared to what was seen in the capsids and appear to have more extensive intra-capsomer contacts. The hexamers and pentamers interact through trivalent connections at the icosahedral and quasi-threefold axes. In contrast to the capsids, the hexamers in the procapsids are completely closed while there is a hole of 1 nm in the center of the pentamers. A cross-section of the procapsid reveals that the serrated appearance of the shell results from tilting of the gp47 subunits relative to the surface of the shell (Fig 11e). The central section also shows a diffuse central density with a diameter of 36 nm that is assumed to represent the scaffolding core. The capsid and core are separated by a 2.5 -nm gap (Fig 11e), suggesting that either there is little direct interaction between the two or that the scaffolding protein is disordered in this region.

SaPI1 capsids

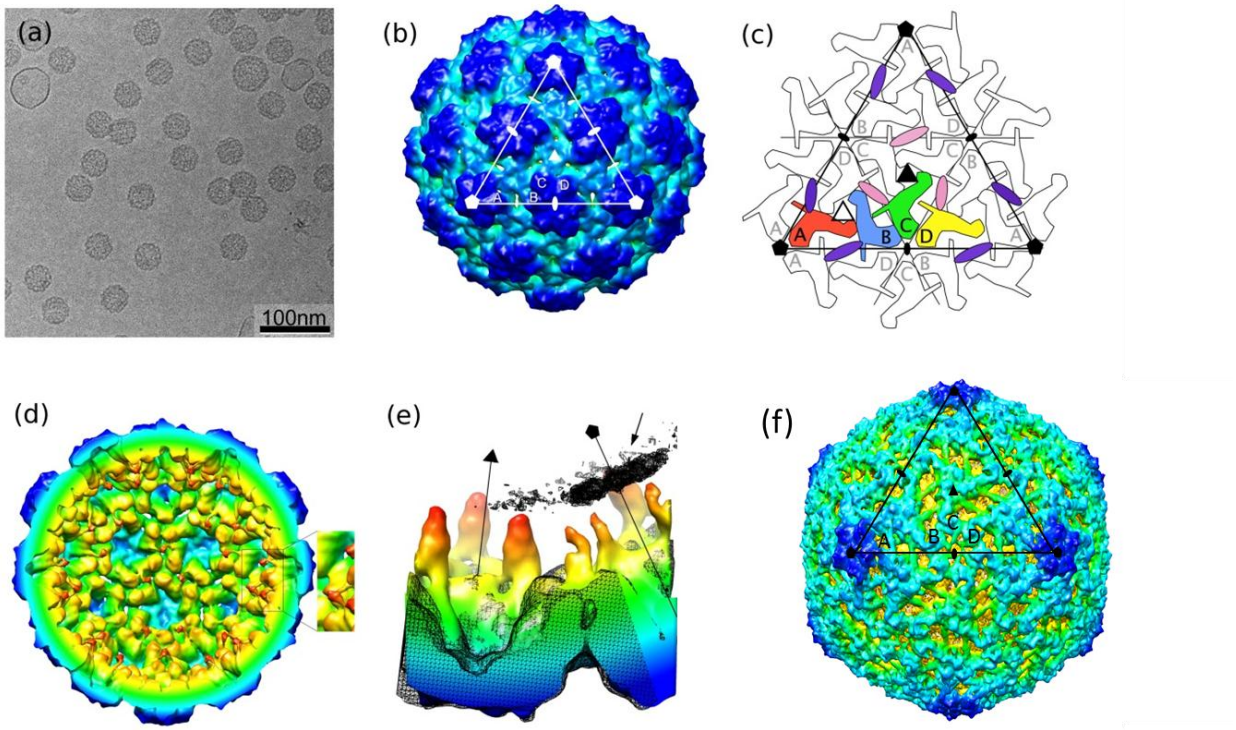
SaPI1 virions have capsids with a 47 nm-diameter and tails similar to those of 80α particles in length and morphology. The interior of the capsids has a punctuate appearance due to the closely packed DNA inside. Symmetric 3-D reconstruction of SaPI1 capsids to a resolution of 1.02 nm shows that SaPI1 has a *T=4* icosahedral symmetry with 240 copies of 80α gp47 organized into 30 hexamers that are located on two adjacent flat triangular faces of the icosahedron in such a way that the sixfold axis of the hexamer coincides with the twofold axis of the icosahedron, and 12 pentamers located on the icosahedral fivefold vertices. (Fig 13). The icosahedral symmetry of T=*4* suggests that gp47 exists in four quasi-equivalent environments as shown schematically in Fig 13. The hexamers and pentamers interact via two types of trivalent connections, a type 1 (CCC) contact at the icosahedral twofold axis, and a type 2 (ABD) contact formed between one pentamer and two hexamers (Fig. 13). As in 80α mature capsids, density in mature SaPI particles is also clustered in distinct trefoils delineated by holes at the contact points. The cross-section of the capsid shows that the shell is on average 2.4 nm thick and has a flat appearance except at the fivefold vertices

SaPI1 procapsids

SaPI1 procapsids were produced from *ΔterS*$_{SaPI1}$ mutant 80α lysogenic strain ST37 and later on from *ΔterS*$_{SaPI1}$ mutant 80α *ΔterS* lysogenic strain ST63. The particles were isolated by procedures described above.

Figure 13. The SaPI1 procapsid.

(a) Cryo-electron micrograph of SaPI1 procapsids (ST63),showing a typical field of particles used for reconstruction. Scale bar, 100 nm. (b) Outside view of the SaPI1 procapsid in isosurface representation, colored radially from the center (red to blue). The large triangle represents one icosahedral face, with the fivefold, threefold and twofold symmetry axes indicated (pentagons, triangle and ovals, respectively). Subunits A, B, C and D of one asymmetric unit are labeled. (c) Schematic diagram showing the triangular face from (b).The four subunits (A, B, C and D) in one asymmetric unit are colored red, blue, green and yellow, respectively. Type 1 and 2 trimers are indicated by the filled and open triangles, respectively. The large ovals represent the threefold-associated (pink) and fivefold-associated (purple) gp6 dimers on the inside of the shell. (d) Inner surface of procapsid, colored radially from the center (red to blue). (e) A magnified view of a portion of the inside of a procapsid (f) Outside view of SaPI1 capsid in isosurface representation, colored radially from the center (red to blue). The large triangle represents one icosahedral face with the fivefold, three fold and twofold symmetry axes indicated by pentagons, triangles and ovals, respectively. The subunits of one asymmetric unit are labeled A, B, C and D.


(a)
100nm
(b)
A
B
C
D
(c)
A
B
C
D
(d)
(e)
(f)
A
B
C
D

SDS-PAGE analysis of sucrose fractions showed a large amount of scaffold and major capsid proteins in fraction F1, indicating that either these proteins remained unassembled or that SaPI1 procapsids are relatively unstable. Fractions F2 and F3, like 80α procapsid fractions, contained predominantly tail proteins and fractions F4 to F6 contained scaffolding and major capsid proteins. The ratio of scaffolding to major capsid protein was approximately 0.2 which is lower than the ratio observed in 80α procapsids. More than half of the particles had expanded, thin walled shells and lacked a scaffold core, suggesting that either SaPI1 procapsids are unstable and liable to spontaneous expansion or that mature SaPI1 capsids that had lost their DNA during the CsCl gradient centrifugation had a similar sedimentation coefficient as SaPI1 procapsids and hence had co-sedimented with them. SaPI1 particles generated from strain ST37 also yielded a higher density band during equilibrium sedimentation in CsCl. Cryo-EM examination of particles from this band revealed that most of them are smaller sized capsids filled predominantly with DNA and about 5% of them were normal sized mature capsid particles. Thus some of the empty expanded shells observed in the sucrose gradient purified fraction could be attributed to loss of DNA.

Symmetric 3-D reconstruction of SaPI1 procapsids to a resolution of 1.0 nm shows that they have a spherical morphology with a diameter of 39 nm and with *T*=4 icosahedral symmetry. There are 240 copies of gp47 in the capsid arranged as pentamers and hexamers. These 240 copies are found in four different environments, termed A, B, C and D. The capsid protein is thus clustered into twelve A5 pentamers located at the fivefold vertices and thirty (BCD)2 hexamers that are located on the icosahedral twofold axes in such a way that the sixfold axis of symmetry of the

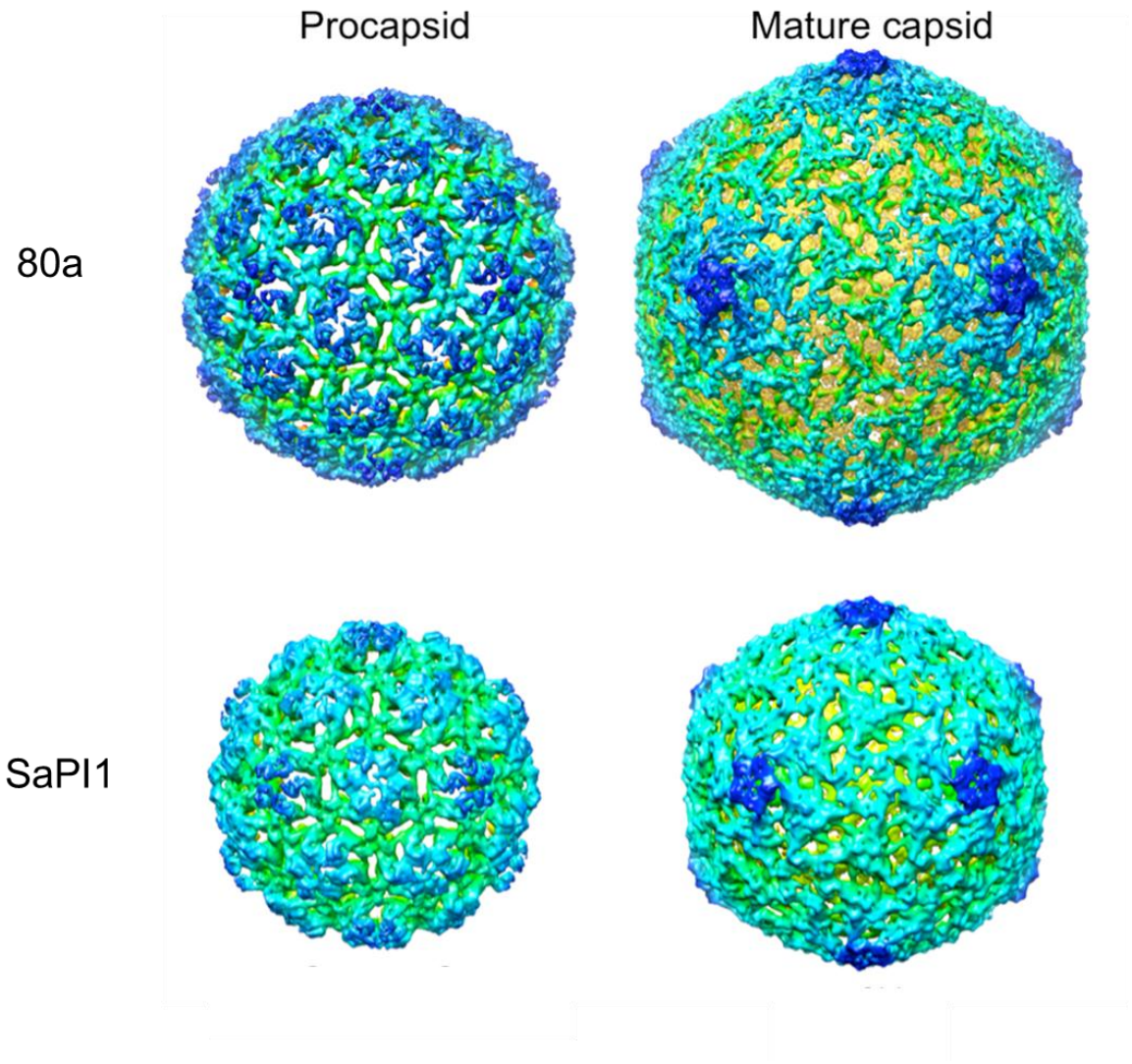
hexamers coincides with the twofold axis of the icosahedron. The hexamers and pentamers interact via two types of trivalent connections, a type 1 (CCC) contact at the icosahedral threefold axis, and a type 2 (ABD) contact formed between one pentamer and two hexamers. There are no additional features that could be attributed to a Sid-like external scaffold on the exterior shell of the SaPI1 procapsid similar to those seen on P4 procapsids. The interior of SaPI1 procapsids however, unlike the 80α procapsids, has 120 nm long finger-like densities that protrude from the inner surface to the center of the procapsid. The density of the finger-like projections is well defined, indicating that the protein is well ordered.

80α major capsid protein has an HK97 like fold

Crystallographic studies with an extensively studied *E. coli* phage, HK97, revealed a unique fold in its major capsid protein, gp5 (Gan *et al.*, 2006). This fold has been subsequently found in all *Caudovirales* studied thus far, including P22 and ε15. Studies with HK97 procapsids have revealed that major conformational changes in the major capsid protein occur during maturation of procapsids, resulting in thin, angularized and expanded shells of the mature virions. HK97 gp5 was found as a close match to the 80α capsid protein based on secondary structure similarity as shown in Fig. 15. Therefore, we modeled 80α gp47 using HK97 gp5 as a template and fitted the model on our cryo-EMs. Further refining was done to obtain the optimal conformation for gp47.

.

Figure 14. Comparison of Icosahedral reconstructions of 80α and SaPI1 capsids and procapsids.
The top two panels show reconstruction images for 80α procapsid and capsid respectively. The procapsid structure was obtained by imaging particles isolated from an 80α *terS* mutant strain ST24. The bottom panel shows reconstruction images of SaPI1 procapsid and capsid. SaPI1 particles were obtained from a SaPI1 positive 80α orf44 mutant strain ST65. The procapsids were obtained from $terS_{SaPI1}$ and $terS_{80\alpha}$ double mutant strain ST63.


Procapsid
Mature capsid
80a
SaPI1

Molecular modeling of gp47 was carried out using two different programs namely, I-TASSER and Genthreader (Jones, 1999; Zhang, 2009). Both programs yielded similar results using HK97 gp5 as a template

The salient features of the structural model of gp47, based on HK97 gp5 illustrated in Fig. 16, are as follows: The HK97-like fold is a unique fold consisting of mixed α/β structure and organized into two compact, spatially distinct domains that are not sequence contiguous, namely the A and P domains, and two extensions, namely the N-arm and E-loop. The axial domain (A domain) is near the fivefold and quasi sixfold symmetry axes, and the peripheral domain (P domain) along with extensions (N-arm and E-loop), occupy the region between adjacent quasi-threefold axes. The domains A and P together give this fold an L-shape. An additional feature not seen in HK97 gp5 is the presence of a P-loop that consists of 13 residues and inserts into the β sheets of the P domain.

A domain

The A domain includes residues 163–254 and 301–324 of gp47 and features two α helices (α5 and α6) and a three β sheets consisting of β3, β6, and β11 that could be confidently fitted into corresponding density in the map. β11 is generated by the C-terminal arm. The remainder of the A domain could not be assigned any secondary structure and is modeled as loops due to either flexibility or limited resolution of reconstruction.

Figure. 15. Sequence and secondary structure of 80α gp47 and HK97 gp5.

The alignment was made by I-TASSER and edited according to the final model. Filled boxes show sequence identity, while open boxes show residues that share physicochemical properties, as defined in the ESPript program. The gp47 sequence starts at Ala15, after the maturation cleavage site. The sequences corresponding to the N-arm, E loop, P domain, and A domain are indicated. The secondary-structure elements of gp47 according to the final model are shown as spirals (α helices) and arrows (β strands) above the alignment. The numbering of the α helices follows that of HK97. The asterisk indicates the kink-forming Pro132 residue in α3.

Reprinted from J Mol Biol. 405 (3) Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids., 863-876, with permission from Elsevier

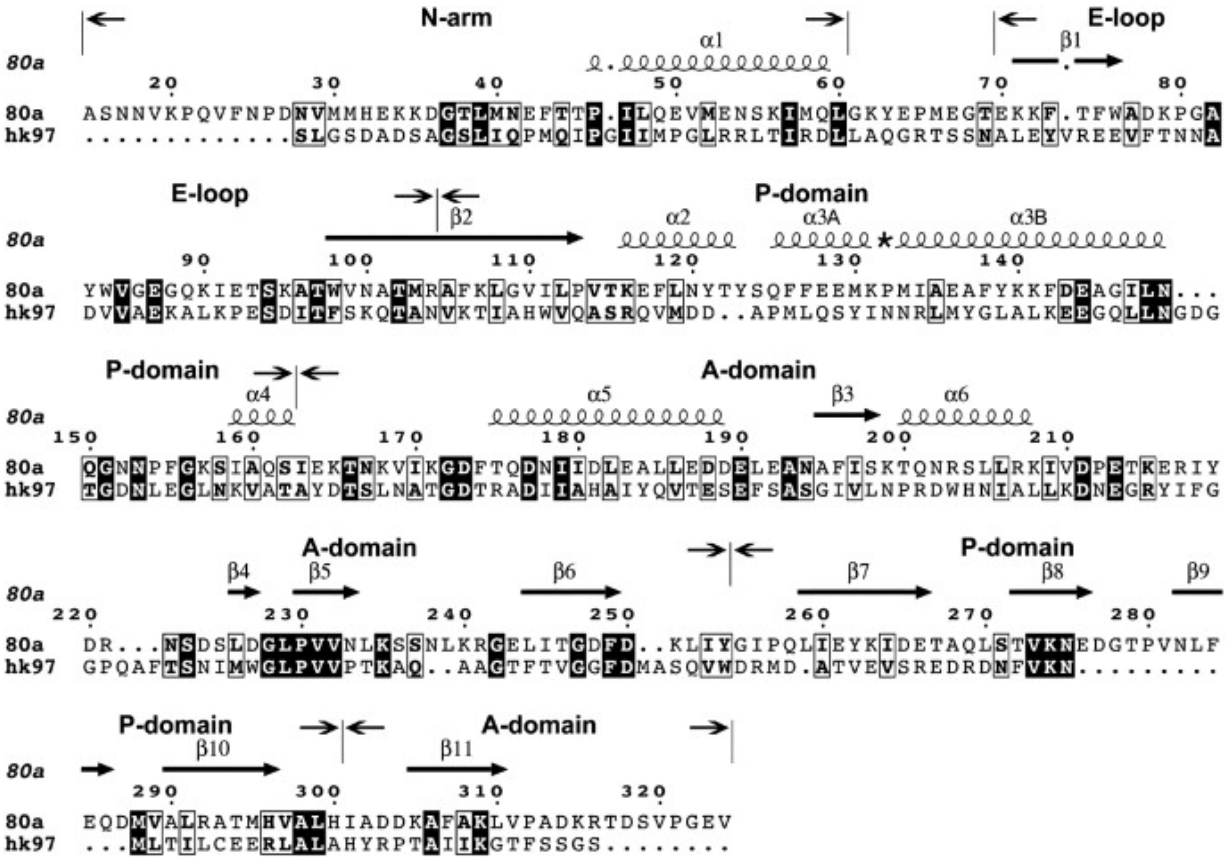
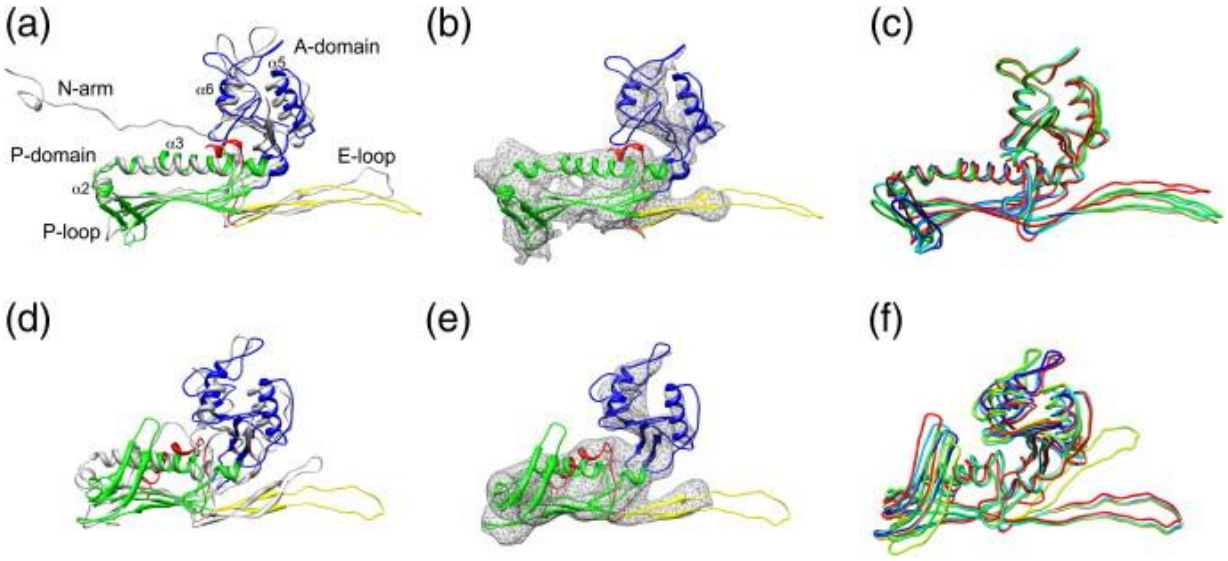

N-arm
E-loop
α1
β1
80a
20 30 40 50 60 70 80
80a ASNNVKPQVFNPDNVMMHEKKDGTLMNEFTTP.ILQEVMENSKIMQLGKYEPMEGTEKKF.TFWADKPGA
hk97 .....SLGSDADSAGSLIQPMQIPGIIMPGLRRLTIRDLLAQGRTSSNALEYVREEVFTNNA
E-loop
β2
P-domain
α2
α3A
α3B
80a
90 100 110 120 130 140
80a YWVGEGQKIETSKATWVNATMRAFKLGVILPVTKEFLNYTYSQFFEEMKPMIAEAFYKKFDEAGILN...
hk97 DVVAEKALKPESDITFSKQTANVKTIAHWVQASRQVMDD..APMLQSYINNRLMYGLALKEEGQLLNGDG
P-domain
α4
A-domain
α5
β3
α6
80a
150 160 170 180 190 200 210
80a QGNNPFGKSIAQSIEKTNKVIKGDFTQDNIIDLEALLEDDELEANAFISKTQNRSLLRKIVDPETKERIY
hk97 TGDNLEGLNKVATAYDTSLNATGDTRADIIAHAIYQVTESEFSASGIVLNPRDWHNIALLKDNEGRYIFG
A-domain
β4
β5
β6
P-domain
β7
β8
β9
80a
220 230 240 250 260 270 280
80a DR...NSDSLDGLPVVNLKSSNLKRGELITGDFD..KLIYGIPQLIEYKIDETAQLSTVKNEDGTPVNLF
hk97 GPQAFTSNIMWGLPVVPTKAQ..AAGTFTVGGFDMASQVWDRMD.ATVEVSREDRDNFVKN.........
P-domain
β10
A-domain
β11
80a
290 300 310 320
80a EQDMVALRATMHVALHIADDKAFAKLVPADKRTDSVPGEV
hk97 ...MLTILCEERLALAHYRPTAIIKGTFSSGS.........

Figure 16. Modeling of gp47.

(a) Ribbon representation of the gp47 (subunit B) model in the 80α virion, colored according to structural domains (red, N-arm; yellow, E loop; green, P domain; blue, A domain) and superimposed on HK97 gp5 (PDB ID: 1OHG) in gray. (b) Ribbon diagram of gp47 fitted into corresponding density from the 80α virion reconstruction. (c) Superposition of the models for the seven nonequivalent subunits in the 80α virion, colored as in Fig 11a. (d) Ribbon representation of the gp47 (subunit B) procapsid model, superimposed on the HK97 gp5 procapsid structure (PDB ID: 3E8K) and colored as in (a). (e) The gp47 model fitted into the 80α procapsid reconstruction. (f) Superposition of the seven nonequivalent subunits in the procapsid.

Reprinted from J Mol Biol. 405 (3) Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids., 863-876, with permission from Elsevier


(a)
A-domain
α5
α6
N-arm
α3
P-domain
E-loop
α2
P-loop
(b)
(c)
(d)
(e)
(f)

P domain

The P domain consists of a prominent spine helix (α3) composed of residues 127–150 of gp47, a shorter α2 helix composed of residues 117–124 and three stranded β sheets namely β2, β7 and β10. β2 is contiguous with the E-loop while the two strands β7 and β10 follow reentry of the sequence into the P domain from the A domain The spine helix α3 has a proline residue at position 132, which is not present in HK97, P22, or ε15, causing the spine helix to be kinked at an angle of 53°. The spine helices in these other bacteriophages are comparatively straight. There is a loop immediately following the spine helix that includes a short helix (α4) and that acts as a hinge between the A domain and the P domain.

P loop

An additional loop not originally found in the HK97 gp5 structure is a 12 residue hairpin loop composed of residues 276–287 that inserts into the β-sheets of the P domain.

E loop

The E loop in gp47 is made up of β1 and β2 sheets composed of residues 70–104 and is of the same length as in HK97.

N-arm

The N-arm consists of residues 15–60 of gp47 and corresponds to the extended N-terminal “N-arm” (residues 104–132) of gp5. In HK97, this arm extends from the body

of the P domain and makes twofold contacts with the corresponding arm from an adjacent capsomer. However, in the 80α virion reconstruction, there was no density corresponding to this sequence

Discussion

3D reconstructions from the cryoelectron images of 80α and SaPI1 capsids and procapsids confirm that they have icosahedral symmetries corresponding to triangulation numbers T=*7l* and T=4, respectively, which is also reflected in the difference in their sizes. The 80α shell is composed of 420 subunits while the SaPI1 shell is made of 240 subunits. Moreover, the major capsid protein exists in seven quasi-equivalent conformations in the 80α capsid and in four different quasi-equivalent states in SaPI1 particles, pointing towards the conformational flexibility of this protein. These seven nonequivalent states are reflected in the differences in the conformational states of gp47 when flexibly fitted into an 80α cryo-electron density map (Spilman *et al.*, 2011).

Since the HK97-like capsid fold has been found in all *Caudovirales* that have been studied thus far, and even in herpes virus (Bamford *et al.*, 2005; Johnson *et al.*, 2007), it is not surprising that the 80α gp47 capsid protein also conforms to the HK97 fold. There are, however, major differences between gp47 and HK97 gp5 throughout the structure. The biggest difference is in the elongated P loop of gp47, which was found to form a trifoliate β hairpin at the threefold inter-capsomer contacts in the 80α icosahedral shell (Spilman *et al.*, 2011). This trifoliate density is reflected in a clustering of strong density in the reconstruction, suggesting that this region is critical for capsid stability. This region has been found to be critical in providing capsid stability in phage λ and P22 as well. For example, in phage λ there is an additional decoration protein (gpD)

at this location that strengthens the three fold intercapsomer interaction (Lander *et al.*, 2008). P22 has an extended P-loop similar to that of gp47 and an additional telokine domain that is involved in capsid stability (Parent *et al.*, 2010). In HK97, residues involved in crosslinking are located in this region (Wikoff *et al.*, 2000). The other important observation is the presence of an α helix in the procapsid that is not found in the mature capsid. This suggests that this arm is involved in the maturation process and possibly interacts with the scaffold protein (Spilman *et al.*, 2011).

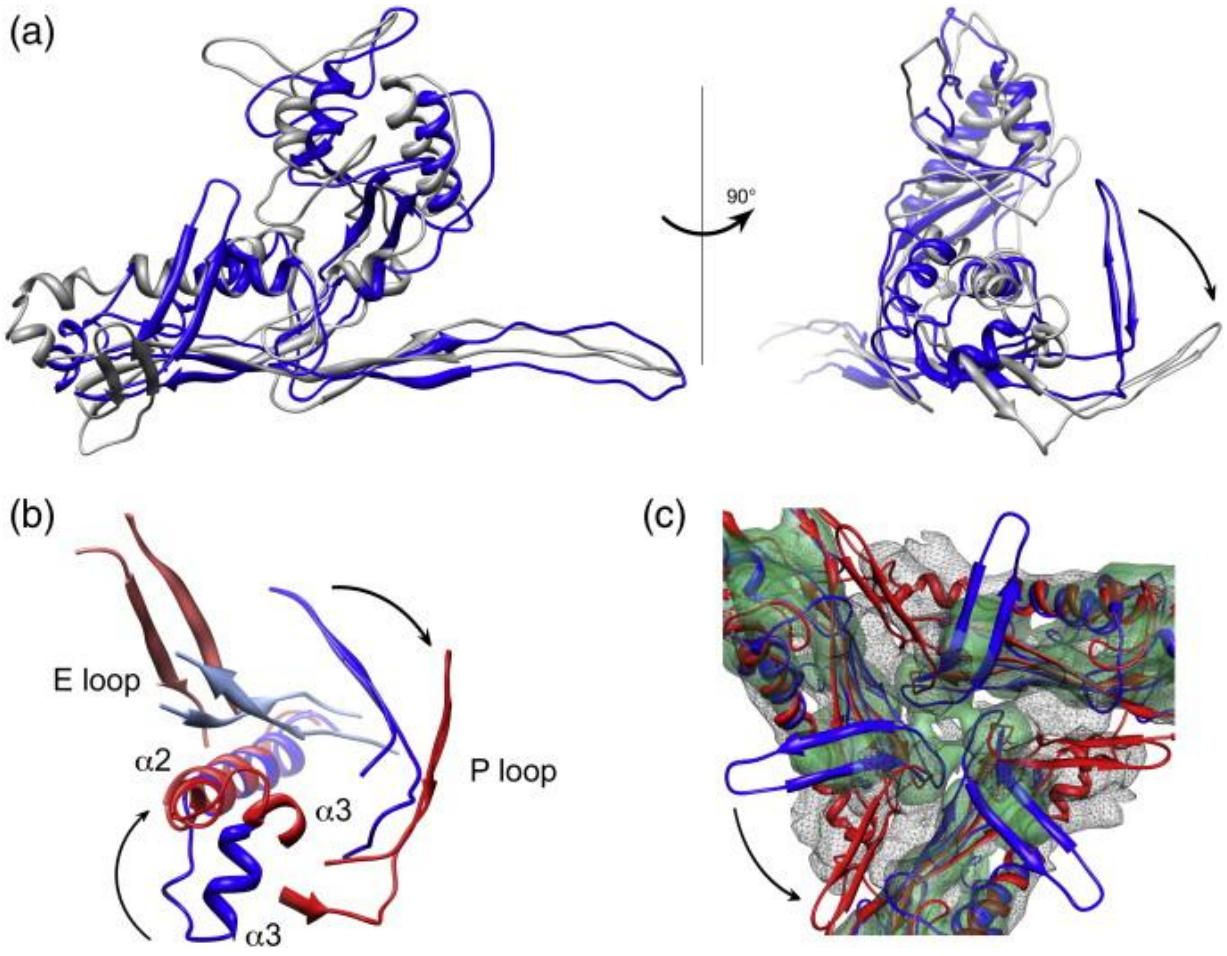
The most striking conformational change in the 80α capsid protein during capsid maturation is in the spine helix (α3), which is composed of residues 127-150 and contains a proline at 132 that produces a kink in the helix. The presence of proline in helices is often associated with conformational changes due to changes in the environment and is used by many biomolecules (Sansom *et al.*, 2000). The flexible fitting of gp47 into procapsid and capsid cryo-electron density maps revealed that ψ angle of Pro132 switches from +135° to – 45°. This change in the ψ angle causes a rotation of the N-terminal half of α3 relative to the rest of the helix (Fig 17a and b) and this change is propagated through α2 to the whole P domain, including the P loop. However, the position of the A domain relative to the P domain remains essentially the same. This mechanism of causing a conformational change differs from HK97, where capsid maturation is caused by a rotation of the A domain by 39° relative to the P domain. Further refinements in this model by improving 3-D reconstructions to better resolution and comparative molecular modeling with SaPI1 particles would provide

valuable insight into the assembly process and help understand the phenomenon of capsid size redirection.

The exterior surface of SaPI1 procapsids surprisingly does not show the presence of any extraneous protein densities resembling Sid protein seen on P4 procapsids. This observation suggests that SaPI1 procapsid assembly is not mediated by an external scaffold like the P4 capsids. However, both 80α and SaPI1 procapsids have internal radial densities which are strikingly prominent in SaPI1 procapsids. The presence of prominent fingerlike projections from the inner surface in SaPI1 procapsids, coupled with the fact that two SaPI1 encoded proteins are present only in procapsids, suggests that these prominent internal densities could be due to SaPi1 proteins. All available data are consistent with an internal scaffold. However, it is also formally possible that interaction of an external scaffold with the external shell is so weak that the external scaffold is lost during the purification steps.

Figure 17. Comparison of the structure of gp47 in the 80α procapsid and virion.

(a) Superposition of the gp47 (subunit B) models in the procapsid (blue) and the virion (gray), aligned on the C-terminal part of the spine helix (residues 133–149). The right panel shows the same superposition rotated by 90° relative to the left panel. (b) Details of interactions between the E loop and the P domain of an adjacent subunit in the procapsid (blue) and virion (red), aligned as in (a). (c) Superposition of the density for the type 1 trimer in the virion (solid, transparent green) and procapsid (mesh) with the corresponding models of gp47 (subunit E) fitted in (blue, procapsid; red, virion). The arrows indicate domain movements upon capsid expansion.
Reprinted from J Mol Biol. 405 (3) Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids., 863-876, with permission from Elsevier


(a)
90°
(b)
E loop
α2
α3
α3
P loop
(c)

Chapter 4

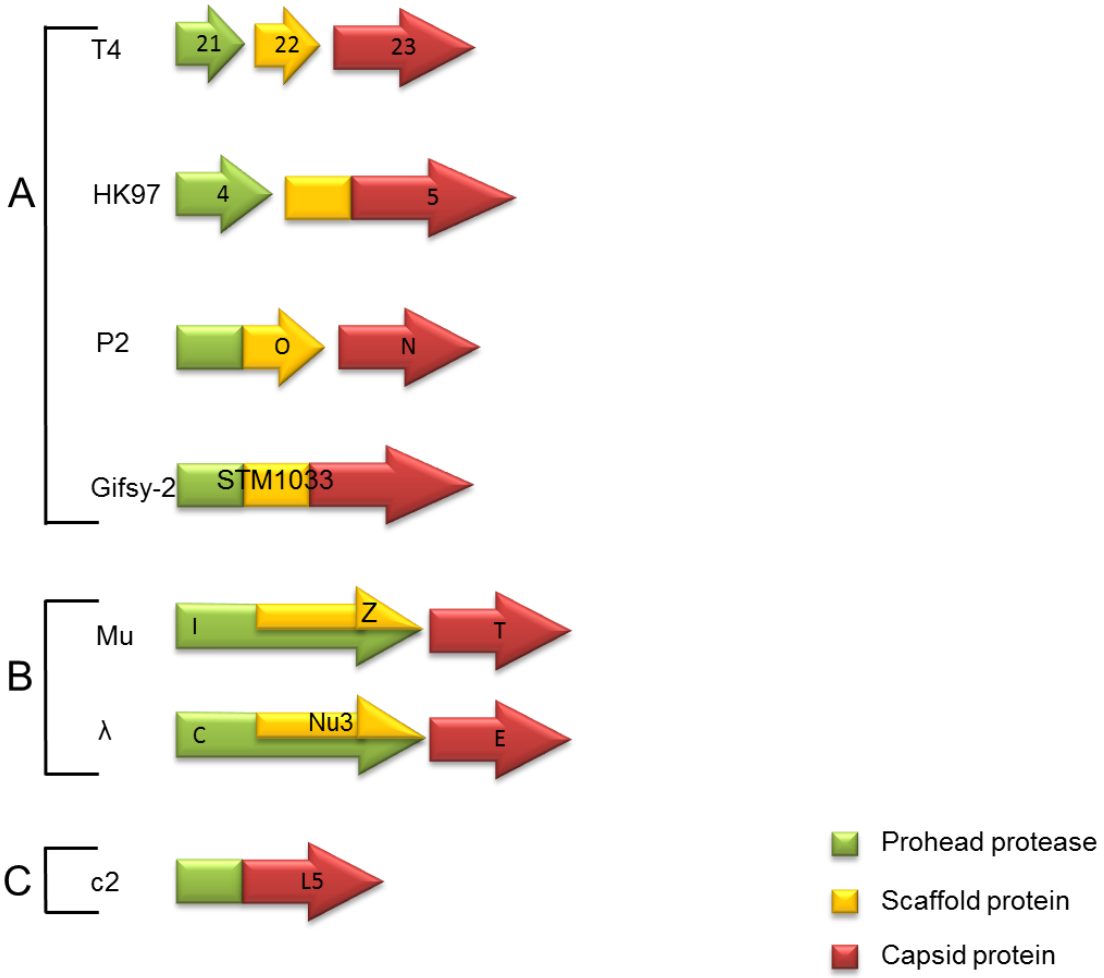
N-terminal cleavage of scaffold and major capsid proteins

Introduction

In many phages, capsid maturation is controlled by a phage protease that is incorporated into the procapsid as a proenzyme that undergoes self-activation. The main role of the phage protease is not degradation of structural components of the procapsid, but to trigger conformational changes by selective cleavage at specific sites in the major capsid protein (Steven *et al.*, 2005). In these phages, the gene encoding the phage protease is a member of the head assembly gene cluster that includes genes encoding the portal protein, scaffolding protein and capsid protein(s) (Hendrix,2003)(Hendrix *et al.*, 1998). Three common arrangements of the location of the protease gene relative to those of the scaffolding and capsid protein genes have been observed (Effantin *et al.*, 2010) (Fig 18). The first arrangement is exemplified by T4, in which the protease gene is followed by the scaffold and major capsid genes. Variants on this arrangement fuse one or more of these genes. In HK97 the protease is followed by fused scaffold and capsid genes (the N-terminal part of gp5 is a scaffold) (Huang *et al.*, 2011). Phage P2 is yet another example of a similar gene order, but in this case the protease is fused to scaffold and the protease activity resides in the N-terminal part of

Figure 18. Schematic representation showing arrangement of proteases in various phages.

A) The phage protease is located between portal and scaffold genes. B) Phage protease and scaffold are encoded by nested genes that have different start codons but same stop codon. C) Phage protease is fused to major capsid gene.


A
T4
21
22
23
HK97
4
5
P2
O
N
Gifsy-2
STM1033
B
Mu
I
Z
T
λ
C
Nu3
E
C
c2
L5
Prohead protease
Scaffold protein
Capsid protein

the protein, while the scaffold activity is located in the C-terminal part (Chang *et al.*, 2009). In a recently studied *Salmonella* phage Gifsy-2, a novel arrangement of composite gene has been found in which the N-terminal part codes for a protease, the central domain of unknown function is implicated to have a scaffold function and the C-terminal region codes for a capsid protein (Effantin *et al.*, 2010). In the second arrangement, the prohead protease and scaffolding protein are encoded by nested genes, as exemplified by those in phage Mu. The prohead protease (gpI) and scaffold (gpZ) have different start codons but same stop codon and use the same reading frame (Morgan *et al.*, 2002). This gene arrangement is also shared by phage λ, in which the phage protease (gpC) and scaffold are encoded by nested genes that are in the same reading frame, have alternate start codons but same stop codon (Morgan *et al.*, 2002; Medina *et al.*, 2010). In the last arrangement, the phage protease is fused to the capsid protein, as exemplified in lactococcal phage c2 (Lubbers *et al.*, 1995). However, not all phages conform to these described categories. It is also worth noting that not all phages encode a protease in their head assembly module, because they do not need it for self-assembly of their procapsids. For example, in phages T7 and P22, head assembly does not involve proteolysis (Duda *et al.*, 1995a)(Hendrix *et al.*, 1998).

Rationale

Previous studies in our lab with peptide mass fingerprinting of 80α and SaPI1 particles clearly demonstrated that the protein composition of both these particles is identical (Tallent *et al.*, 2007). Furthermore, structural studies with 80α and SaPI1 mature particles and their precursors confirmed that the particles first assemble as roughly spherical procapsids of two distinct sizes that undergo morphological changes

to yield mature particles that are expanded, have thinner walls and icosahedral symmetry corresponding to two different triangulation numbers (Chapter 3). During our attempts at *in vivo* procapsid assembly in *E. coli* involving over-expressed scaffold and major capsid proteins, we discovered that the masses of scaffold and major capsid proteins expressed in *E. coli* were higher than those of the native ones found in virions. The most likely explanation for this discrepancy in masses was that the proteins were being processed during capsid assembly. In order to investigate how these proteins were post-translationally modified, we analyzed them by SDS-PAGE and determined their masses by ESI-mass spectrometry.

Results

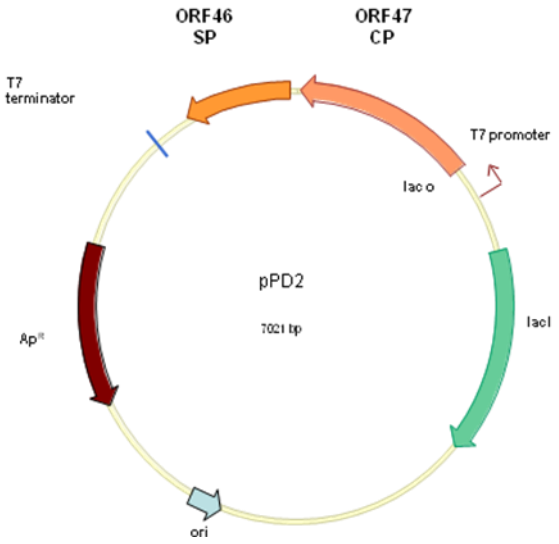
80α major capsid and scaffolding proteins are post-translationally cleaved at their N-termini

Co-expression of 80α scaffold and major capsid proteins in *E. coli* was carried out from plasmid pPD2, a derivative of pET21a, by induction with IPTG (Poliakov *et al.*, 2008) (Fig 19a). The cells were harvested, lysed and the proteins were analyzed by SDS-PAGE along with proteins from purified native 80α and SaPI1 procapsids. Furthermore, the co-expressed proteins were purified by ultracentrifugation and examined under an electron microscope for formation of procapsid like particles. Comparison of over-expressed scaffold and major capsid with those from native procapsids by SDS-PAGE reveals a clear discrepancy in migration rates (Fig 19b). Both gp46 and gp47 from native procapsids migrated faster than the over-expressed ones by

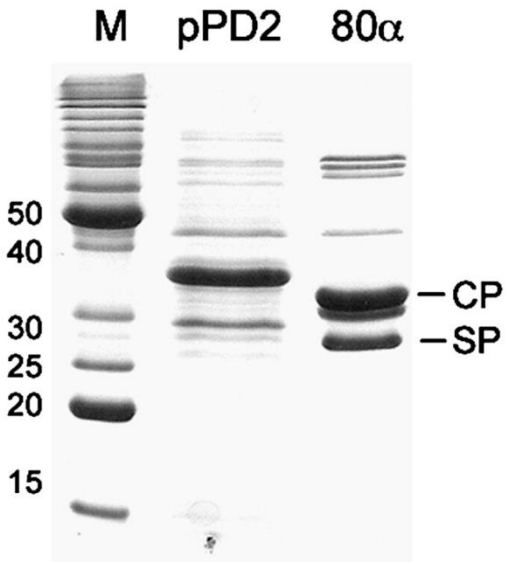
Figure 19. Effect of expressing major capsid and scaffold proteins in *E. coli*

(a) Map of plasmid pPD2 used for overexpression of gp46 and gp47 *in E. coli* (b)SDS-PAGE of major capsid and scaffolding proteins expressed from pPD2 in *E. coli* compared to the same proteins from 80α procapsids. The bands corresponding to gp46 (SP) and gp47 (CP) are indicated on the 80α lane. (c) Electron micrograph of amorphous structures formed during co-expression of major capsid and scaffold proteins in *E. coli.*

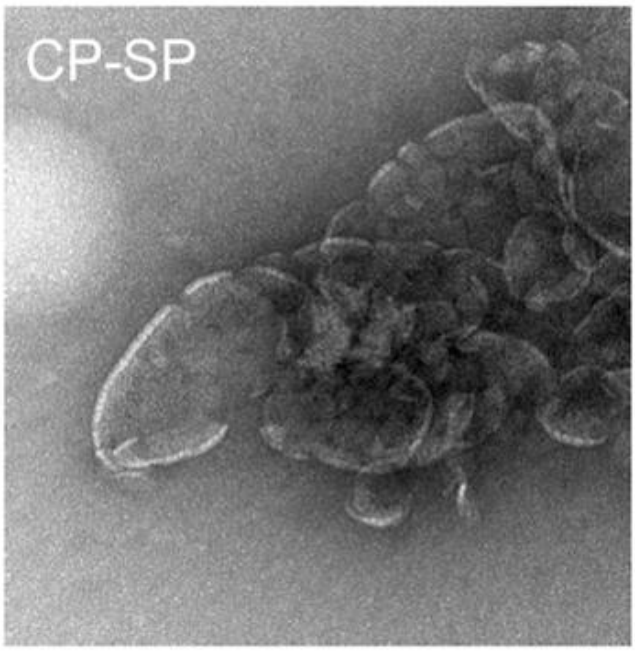
A


ORF46
SP
ORF47
CP
T7
terminator
T7 promoter
lac o
pPD2
7021 bp
lacI
Ap^R
ori

B


M
pPD2
80α
50
40
30
25
20
15
CP
SP

C


CP-SP

at least 2 to 3 kDa, suggesting that scaffold and major capsid are post-translationally cleaved in the native procapsids. Furthermore, self-assembly of co-expressed scaffold and major capsid proteins failed to yield procapsid like particles and instead assembled into monster-like structures similar to those observed for phage P22 (Earnshaw *et al.*, 1978) (Fig 19c) indicating that either these proteins were improperly folded in *E. coli* or proteolytic cleavage of scaffold and major capsid proteins does not occur in this system and is necessary for self-assembly of procapsid like particles.

Our collaborators performed mass spectrometry (LC-ESI-MS/MS) of structural proteins from 80α and SaPI1 procapsids to confirm the above stated observations and determine the molecular masses of major capsid and scaffold proteins (Poliakov *et al.*, 2008). The measured masses of 35,062.20 ± 0.18 Da in 80α and 35,061.86 ± 0.53 Da in SaPI1 are within 1 Da of the calculated mass (35,062.9 Da) for residues 15–324 of the 80α major capsid protein (gp47). The measured masses of 21,700.5 ± 0.07 Da (in 80α) and 21,701.05 ± 0.34 Da (in SaPI1) are within 1 Da of the calculated mass (21,700.9 Da) for residues 14–206 of the 80α scaffolding protein gp46 (Fig 20a).

These data show that scaffold and major capsid proteins are identically processed in both 80α and SaPI1 procapsids and rules out the possibility of differential protein processing in formation of smaller capsids. Furthermore, the most interesting observation is that the scissile bond in both the scaffold and major capsid protein sequences lies between Phe and Ala. Not only are the P1-P1' residues identical, but there is also significant similarity in the N-terminal region (P9-P1) of the major capsid

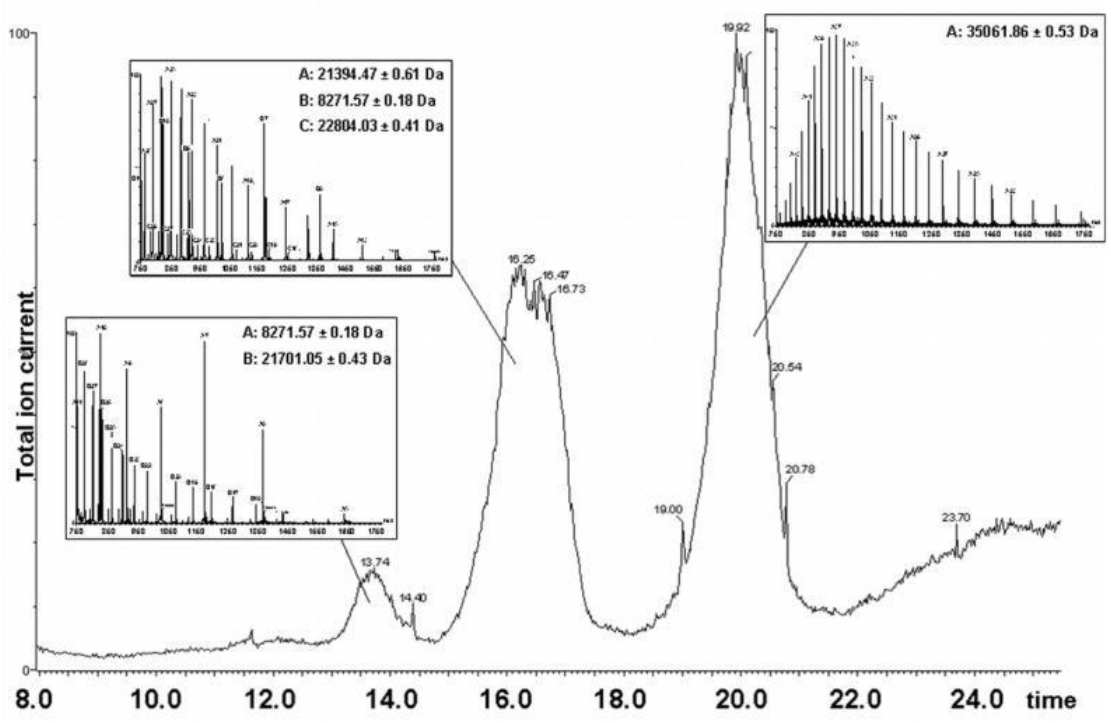
Figure 20. Mass spec analysis and cleavage sites of major capsid and scaffold proteins

Reverse phase chromatography and ESI-MS of SaPI1 procapsids.

The total ion current of the MS detector is plotted against elution time in minutes. The ion current does not accurately reflect the abundance of the measured ionic species. Peaks and their corresponding spectra are shown with the measured masses listed. Peak one (approximately 14 min) includes SaPI1 gp6 (theoretical mass 8,271.5 Da) and N-terminally truncated scaffolding protein gp46 (21,700.9 Da); peak two also contains gp6 as well as SaPI1 gp7 (22,805.3 Da) and major tail protein gp53 (21,394.5 Da). The third peak in the chromatogram contains N-terminally truncated major capsid protein gp47 (35,062.9 Da).
(B) Table of measured masses of gp46 and gp47 in 80α and SaPI1 procapsids, compared to the theoretical mass for the cleaved protein, as well as the N-terminal sequence of the two proteins. The cleavage site is indicated by an asterisk (*)

Reprinted from J Mol Biol. 380 (3) Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids, 465-475, (2008) with permission from Elsevier

A



B

	Measured mass (Da)		Theoretical mass (Da)	Sequence (* = cleavage site)
	80a	SaPI1		
SP, gp46	21,700.5	21,701.1	21,700.9	MEENKLKFNLQFF*ADQSDDPDEPGGDGKKGNPDKKEND...
CP, gp47	35,062.2	35,061.9	35,062.9	MEQTQKLKLNLQHF*ASNNVKPQVFNPDNVMMHEKKDGT...

and scaffolding proteins with a consensus sequence of KLKxNLQxF*A, where * denotes the cleavage site (Fig 20b). All these results point towards involvement of a common protease in N-terminal cleavage of both the major capsid and scaffold proteins.

Cleavage of structural proteins is not uncommon among dsDNA bacteriophages and has been reported in phages λ (Medina *et al.*, 2010), P2 (Chang *et al.*, 2008), T4 (van Driel *et al.*, 1980) and HK97 (Duda *et al.*, 1995b). Such cleavage has been observed to occur after procapsid assembly in P2 (Marvik *et al.*, 1994a) and HK97 (Conway *et al.*, 1995) and is considered an important control point in the assembly pathway that is essential for capsid expansion to occur. The effect of the cleavage may be to affect the thermodynamic stability of the capsid through changes in quaternary interactions or to induce release of the scaffolding proteins. Cleavage of the capsid protein with trypsin in phage P4 procapsids led to capsid expansion and scaffold removal (Wang *et al.*, 2003). Cleavage and even complete degradation of the scaffolding protein are commonly used mechanisms for scaffolding removal and escape from the procapsid in some viruses (Dokland,1999; Fane *et al.*, 2003). Proteolytic cleavage is usually carried out by a phage-encoded protease. As described above, the protease gene is generally located immediately upstream of the scaffolding protein in the same operon, or sometimes embedded in the scaffolding protein itself. Staphylococcal phage proteases have not been extensively studied. However, in some staphylococcal phage genomes, such as phiN315 and 77, orfs with ClpP like domains belonging to the crotonase-like superfamily (NP_835555.1 and NP_958605.1 respectively) are found immediately upstream of the capsid genes. Thus, it seemed likely that one of three putative orfs that lie between the scaffold and portal genes in the

80α genome, namely orf43, orf44 and orf45, might encode a phage protease. orfs 43 and 45 are both extremely small, encoding putative proteins of 61 and 56 amino acids, respectively, and deletions of the corresponding orfs in the related phage ϕ11 have no phenotype (J. Penades, personal communication). The product of orf44, gp44, was found in both 80α and SaPI1 procapsids (Poliakov *et al.*, 2008), and previously in 80α virions (Tallent *et al.*, 2007) and thus was the most likely candidate for a phage protease.

Role of gp44 in self-assembly of 80α and SaPI1 capsids

In order to determine whether gp44 is the phage protease, we deleted orf44 from an 80α prophage in both SaPI1 negative and SaPI1 containing *S. aureus* strains to generate *S. aureus* strains ST64 and ST65, respectively. We expected that deletion of a gene encoding the phage protease would lead to arrest of the assembly process at the procapsid stage because of uncleaved scaffold and major capsid proteins, and hence would be lethal (Medina *et al.*, 2010). Furthermore, we also expected that the arrest of self-assembly at the procapsid stage would lead to a slower migrating full length major capsid protein and the presence of an extra band of scaffold protein, completely absent in mature particles, during SDS-PAGE analysis. Also, the morphology of the particles obtained from the mutant strains would be procapsid-like and significantly different than those from wt mature particles and could be easily distinguished by electron microscopy.

S. aureus strains ST64 and ST65, along with the 80α wt equivalent strains RN10616 and RN10628 as controls, were induced by mitomycin C and monitored until

lysis. Cell lysates were filtered through 0.45 μm filters and the phage concentration and transducing titers were determined. Both ST64 and ST65 cell lysates yielded no detectable pfu (Fig 21a). Thus, orf44 is essential for the phage. Surprisingly, orf44 is non-essential for the SaPI1 transducing particles; we observed a SaPI1 transduction frequency of 4.5 x 10^8 TU/ml, comparable to wt SaPI1 particles (Fig 21b).

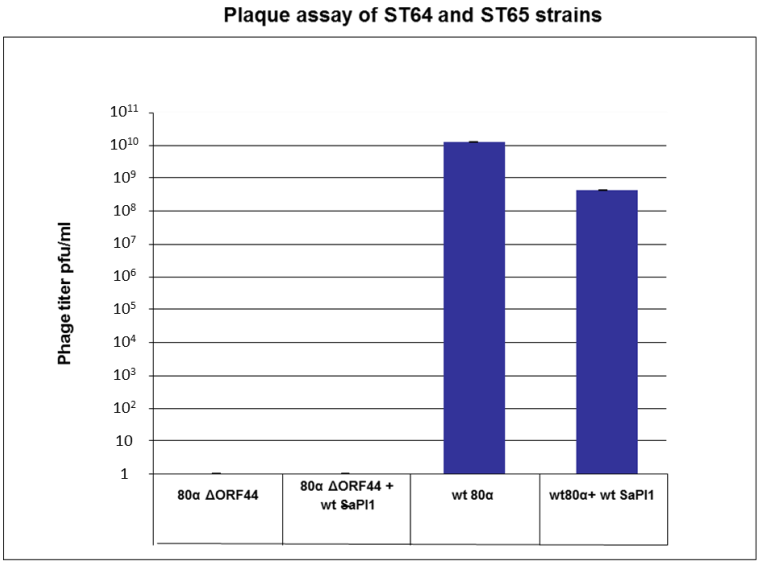
Particles from ST64 and RN10616 lysates were also precipitated using $ZnCl_2$ and subjected to SDS-PAGE analysis in order to compare the migration rate of their major capsid proteins. The migration rates of major capsid from orf44 mutant 80α and wt 80α were found to be the same (Fig 22). Also, there was no extra band that could be attributed to the scaffold. These results suggested that the scaffold was being removed and hence the self-assembly was not arrested at the procapsid stage. Furthermore, in our collaborator's lab, electron microscopy was performed on particles obtained from ST64 and ST65. We observed expanded aberrant capsids with tails attached in electron micrographs of ST64 samples (Fig 23a). This observation supported our view that the mature capsids were being formed. However, we also observed an anomalously large number of free tails and some capsids without tails, suggesting that gp44 had a significant role in capsid stability. Only a few of these expanded capsids had DNA in them. Electron micrographs of SaPI1 containing strain ST65 (Fig 23b), however, showed predominantly normal looking, smaller sized, DNA filled capsids with attached tails, and there were no free tails or empty capsids seen in the micrographs.

Figure 21. Effect of orf44 deletion on 80α and SaPI1 titers.

SaPI1 negative (ST64) and SaPI1 positive (ST65) lysogens of 80α Δorf44 were grown to mid-exponential phase and then induced with 2 µg/ml mitomcyin C. Shown in the panel (a) are the plaque titers and in panel (b), the transducing titers of these strains with RN10616 and RN10628 as controls

.

(a)


Plaque assay of ST64 and ST65 strains
Phage titer pfu/ml
10^11
10^10
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10
1
80α ΔORF44
80α ΔORF44 + wt SaPI1
wt 80α
wt80α+ wt SaPI1

(b)

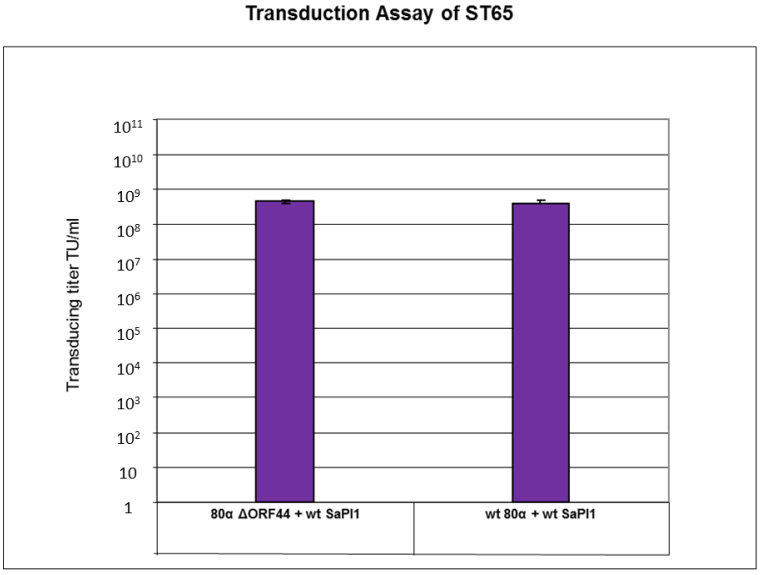

Transduction Assay of ST65
Transducing titer TU/ml
10^11
10^10
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10
1
80α ΔORF44 + wt SaPI1
wt 80α + wt SaPI1

Figure 22. SDS-PAGE analysis of proteins from gp44 mutant phage and SaPI1 particles.

Comparison of major capsid proteins from SaPI1 negative (ST64) and SaPI1 positive (ST65) 80α Δorf44 mutant lysogenic strains were induced with mitomycin C and allowed to lyse. Particles were purified using $ZnCl_2$ and analyzed side by side with those from wt 80α. Electrophoresis was done on 12% Tris-HCl gel.

Lane (a) molecular weight
(b) 80α particles
(c) particles from 80α Δorf44 (ST64)
(d) SaPI1particles with an 80α Δorf44 helper phage (ST65)

The arrow indicates the band corresponding to truncated major capsid protein in each lane.

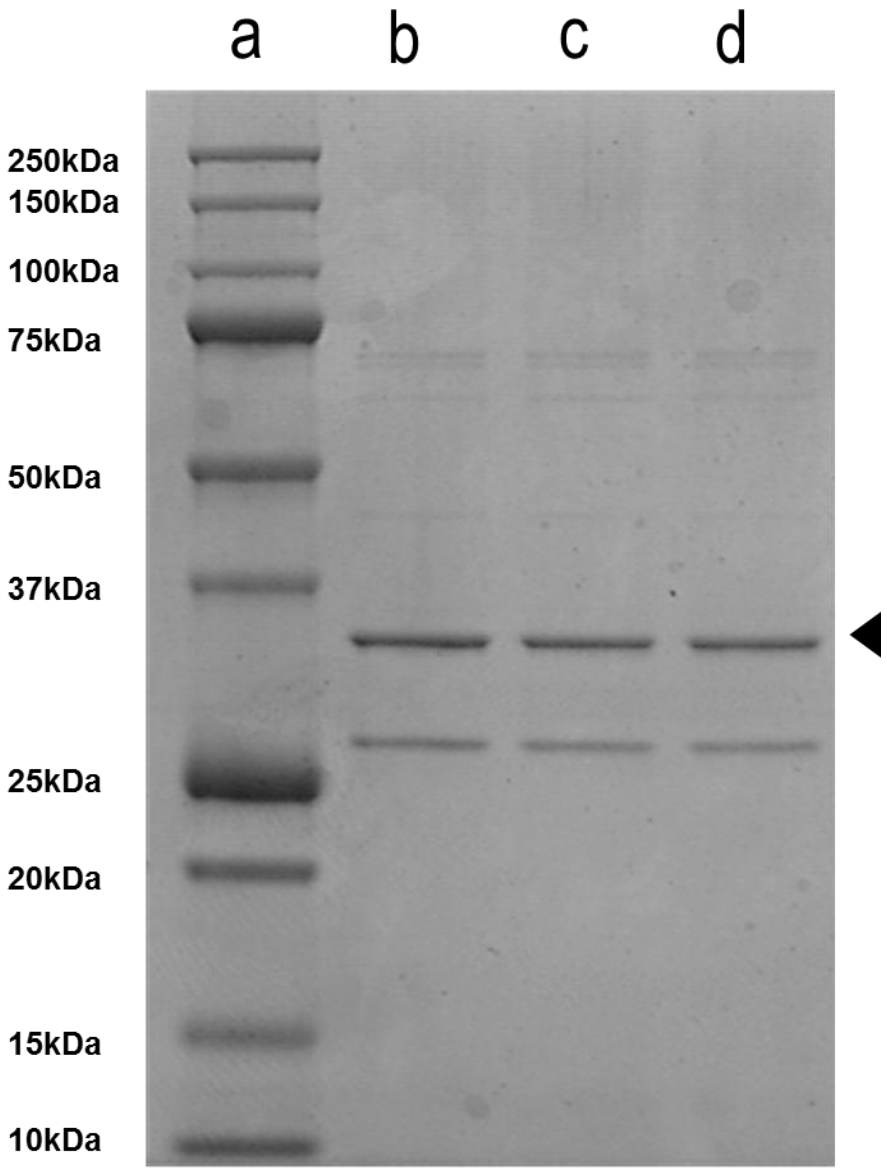

a
b
c
d
250kDa
150kDa
100kDa
75kDa
50kDa
37kDa
25kDa
20kDa
15kDa
10kDa

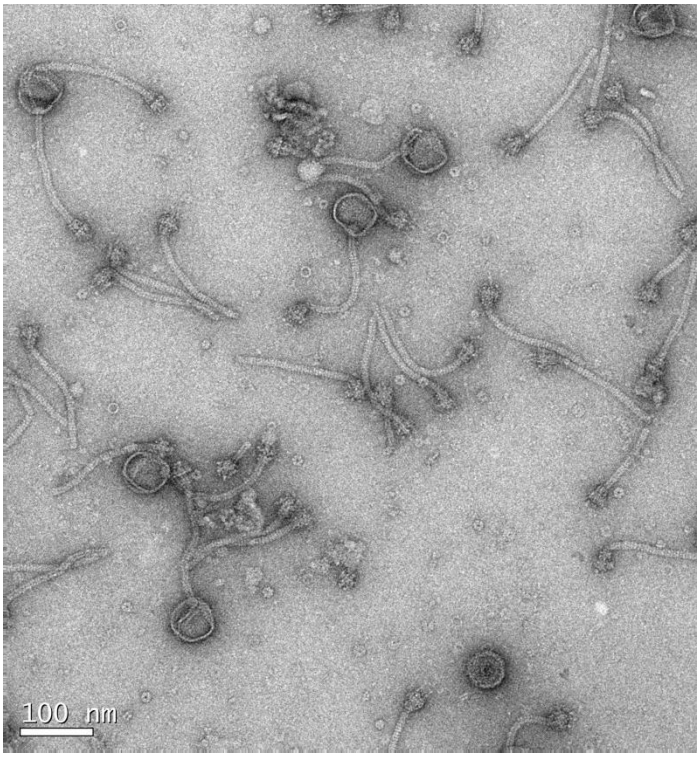
Figure 23. Electron micrographs of CsCl purified phage particles obtained from orf44 deletion mutants.

Samples for negative stain were prepared by applying phage suspensions to glow-discharged carbon-only grids (Electron Microscopy Sciences), washed 2× with dialysis buffer and stained with 1% uranyl acetate. Samples were observed in an FEI Tecnai F20 electron microscope operated at 200kV and images were captured on a 4k × 4k Gatan Ultrascan CCD camera or on Kodak SO-163 film at magnifications from 38,000× to 81,200×.

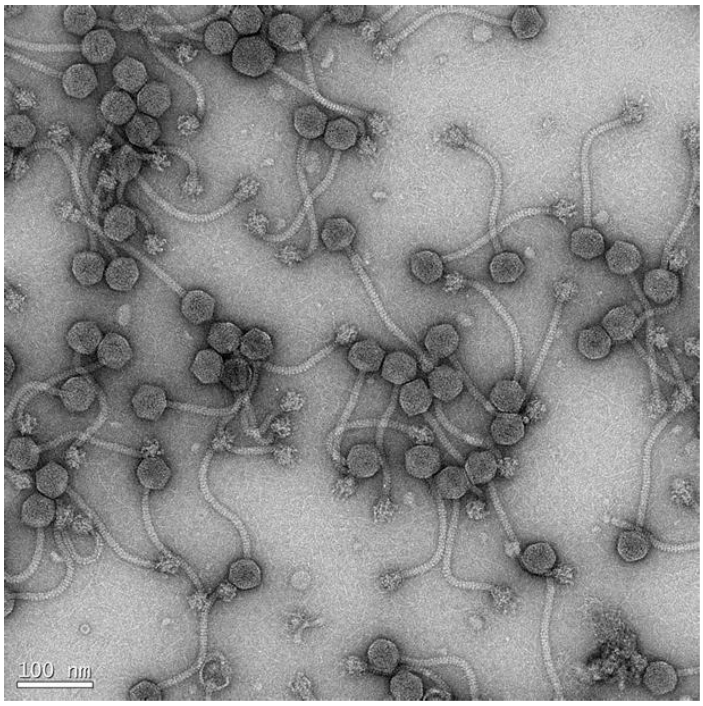
80α∆44 particles from ST64.

SaPI1 particles with an 80α∆44 helper phage, from ST65.

a. 80α



b. SaPI1



SaPI1 capsids are more stable in absence of 80α gp44

Cryo-electron microscopy of wt SaPI1 particles always posed a technical problem. Sample preparation for cryo-electron microscopy of SaPI1 particles was difficult because good quality ice could not be generated with the specimen samples, due to the presence of excessive amounts of free DNA. This resulted in sub-optimal imaging conditions and hence it was difficult to collect a large number of images from a given sample. This created a major setback in 3D cryo-electron reconstruction of SaPI1 capsids because the resolution of a reconstruction is directly proportional to the number of images. The problem of free DNA in SaPI1 lysate is so acute that one can detect it by "streaming" of the sample when collecting CsCl banded SaPI1 particles. The free DNA is believed to have leaked from SaPI1 particles during the process of centrifugation. Interestingly, this free DNA was not found in the CsCl bands during isolation of SaPI1 particles generated from the SaPI1-positive 80α Δgp44 mutant strain. Furthermore, when these samples were prepared for cryo-electron microscopy, they vitrified well. These observations suggest that SaPI1 capsids are more stable in the absence of 80α gp44. Cryo-electron 3D reconstructions of SaPI1 capsids were subsequently carried out using this strain.

gp44 is not a protease

The results from SDS-PAGE analysis clearly ruled out gp44 as the prohead protease, since the migration rate of major capsid protein from the 80α gp44 mutant is same as from wt 80α. In addition, no extra band corresponding to scaffold was observed. This means that major capsid protein is processed and that scaffold is removed from the shells. Results from electron microscopy also clearly indicated that

mature particles were being formed. Furthermore, transduction assays showed that SaPI1 particles capable of transducing at wt levels were being assembled from the mutant 80α in presence of SaPI1. It was surprising that gp44 was essential for 80α, but not for SaPI1. These observations strongly suggested that gp44 is a structural protein and possibly plays a role in stability of T=7 capsids. This suggestion is supported by our observation of an abnormal number of free tails and some filled capsids without tails in the electron micrographs of ST64 samples. Since the samples for electron microscopy were prepared by purifying phage particles by isopycnic sedimentation on a CsCl gradient, a process that separates molecules on the basis of density, it is unusual that the free tails co-purified with intact DNA-containing particles. Thus it seems more likely that these free tails resulted from disruption of intact particles after the purification process and suggests that the stability of the particles is affected by loss of gp44.

A search in the reference protein database (NCBI) using PSI-BLAST (Position-specific iterated BLAST) (Altschul *et al.*, 1997) with the gp44 amino acid sequence as a query showed that this protein had homologs in other staphylococcal *Siphoviridae* and a slight similarity to phage SPP1 gp6 protein. SPP1 gp6 is required for efficient infection of *B.subtilis* and has been demonstrated to play a role in DNA ejection (Vinga *et al.*, 2006). Deletion of the gene encoding SPP1 gp6 causes premature DNA ejection and significant loss of infectivity. Thus based on all these data, we can safely rule out the possibility that orf44 encodes a phage protease and suggest that it is a “neck” protein involved in head-tail connection and/or DNA stability.

A host protease is involved in N-terminal cleavage of major capsid and scaffold proteins.

The conclusion that gp44 is not a phage protease led us to consider other possibilities for the protease activity. Sometimes the protease is found embedded in the scaffold (Conway *et al.*, 1995; Duda *et al.*, 1995a) or the major capsid protein (Morillas *et al.*, 2008) and there were three more orfs in the 80α morphogenetic operon, namely orf43, orf45 and orf48 that did not have any functions assigned to them. The possibility that scaffold or the major capsid protein had proteolytic domains was remote since neither of these proteins was truncated when expressed in *E. coli*. However, proteolytic inactivity due to improper folding could not be ruled out since we had also failed to obtain discrete procapsid like particles from these proteins.

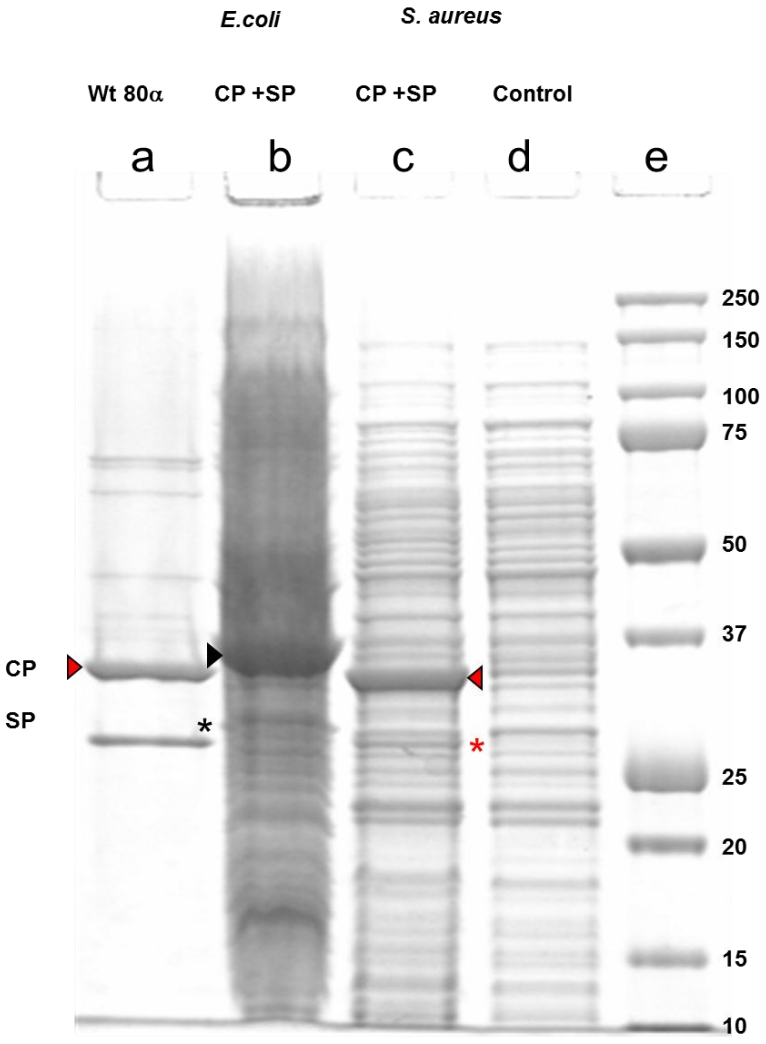
In order to rule out the possibility that proteolytic inactivity was due to improper folding in *E. coli*, we over-expressed scaffold and major capsid proteins in *S. aureus*, the normal host of 80α, by cloning them into shuttle vector pG164 (D'Elia *et al.*, 2006; Pereira *et al.*, 2008) to generate plasmid pPD21 and expressing the genes in strain SA178RI which encodes T7 RNA polymerase. Strains carrying pPD2 and pPD21 were grown and induced with IPTG for 2 hours, after which cells were harvested and the proteins analyzed by SDS-PAGE. This analysis (Fig 24) showed that the migration of both the scaffold and major capsid proteins expressed in *S. aureus* was similar to those of native virions and faster than those expressed from the same genes in *E. coli*. Thus, both scaffold and major capsid proteins were cleaved in *S .aureus* in the absence of any other 80α-encoded genes. This observation ruled out the possibility that any other 80α

Figure 24. Cleavage of 80α scaffold and capsid proteins in *S. aureus.*

SDS-PAGE analysis of overexpressed scaffold and major capsid proteins in *E. coli* and *S. aureus, compared* with those obtained from mature virions.

(a) wt 80α virions
(b) Plasmid pPD2, expressing major capsid and scaffold in *E. coli.*
(c) Plasmid pPD21, expressing major capsid and scaffold in *S. aureus.*
(d) Plasmid pG164 as a negative control.
(e) MW markers; size in KDa indicated

Arrowheads indicate the bands corresponding to capsid (cp) and asterix to scaffold (sp) proteins. The full length proteins are indicated in black and their processed versions in red.


E.coli
S. aureus
Wt 80α
CP +SP
CP +SP
Control
a
b
c
d
e
250
150
100
75
50
37
25
20
15
10
CP
SP
*
*

orfs encoded the phage protease. Furthermore, the mass of major capsid when expressed alone in *S. aureus* was also similar to the native protein from 80α virions (Fig 35), eliminating the possibility that the scaffold is responsible for the proteolytic activity. These observations suggested that either the major capsid protein contains a proteolytic domain or its proteolysis is effected by a host protease.

A BLAST search using the major capsid protein sequence as the query failed to reveal any conserved domains that correspond to known proteases This led us to consider the possibility that the 80α scaffold and capsid were cleaved by a host protease. If indeed the host protease cleaved these phage proteins then there would be host proteins that are targets of this protease. Our earlier studies showed that 80α scaffold and major capsid proteins are N-terminally cleaved and that the scissile bond lies between Phe and Ala, within a conserved sequence in both these proteins, suggesting the possibility of this region being involved in substrate specificity. Therefore it is likely that the targets of this protease would also have similar sequence upstream of their cleavage sites. A BLAST search of the N-terminal cleaved fragment in the *S. aureus* protein database yielded a single hit, *S. aureus* ribosomal protein L27

Ribosomal protein L27 is a protein from the large (50S) ribosomal subunit that participates in both subunit assembly and the peptidyl transferase reaction (Wower *et al.*, 1998). An L27 deletion mutant of *E. coli* shows retarded growth (five to six fold) compared to the wild-type strain, as well as deficiency in the peptidyl transferase activity (Maguire *et al.*, 2005). Crystallographic study of the full ribosome (70S) from *Thermus thermophilus* showed that the N-terminus of the ribosomal protein L27 extends into the

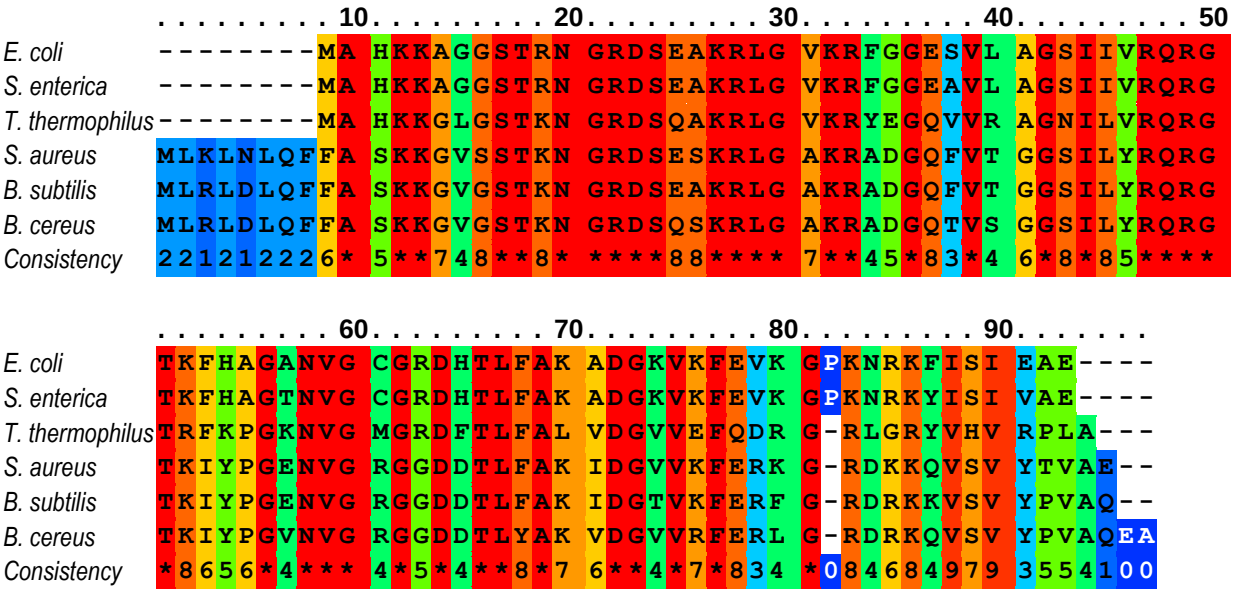
Figure 25. Sequence alignment of ribosomal protein L27

(a) Shown is a sequence alignment from several representative Gram negative and Gram positive bacteria showing the difference in the N-terminal region.
(b) Sequence alignment of the N-terminal sequences of 80α major capsid protein, scaffold protein and S. aureus ribosomal protein, L27

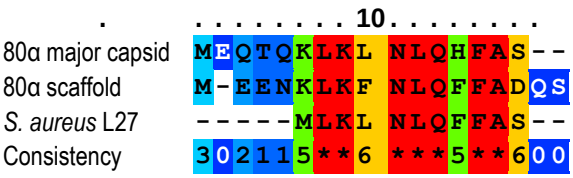
The alignment was performed using an online multiple sequence alignment program: Praline, IBIVU Vrije Universiteit Amsterdam

The conservation scoring is performed by PRALINE. The scoring scheme works from 0 for the least conserved alignment position, up to 10 for the most conserved alignment position.
The color assignments are:

(a)


. . . . . 10 . . . . . 20 . . . . . 30 . . . . . 40 50
E. coli --------MA HKKAGGSTRN GRDSEAKRLG VKRFGGESVL AGSIIVRQRG
S. enterica --------MA HKKAGGSTRN GRDSEAKRLG VKRFGGEAVL AGSIIVRQRG
T. thermophilus --------MA HKKGLGSTKN GRDSQAKRLG VKRYEGQVVR AGNILVRQRG
S. aureus MLKLNLQFFA SKKGVSSTKN GRDSESKRLG AKRADGQFVT GGSILYRQRG
B. subtilis MLRLDLQFFA SKKGVGSTKN GRDSEAKRLG AKRADGQFVT GGSILYRQRG
B. cereus MLRLDLQFFA SKKGVGSTKN GRDSQSKRLG AKRADGQTVS GGSILYRQRG
Consistency 22121222 6* 5**748**8* ****88**** 7**45*83*4 6*8*85****
. . . . . 60 . . . . . 70 . . . . . 80 . . . . . 90 . . . . . .
E. coli TKFHAGANVG CGRDHTLFAK ADGKVKFEVK GPKNRKFISI EAE----
S. enterica TKFHAGTNVG CGRDHTLFAK ADGKVKFEVK GPKNRKYISI VAE----
T. thermophilus TRFKPGKNVG MGRDFTLFAL VDGVVEFQDR G-RLGRYVHV RPLA---
S. aureus TKIYPGENVG RGGDDTLFAK IDGVVKFERK G-RDKKQVSV YTVAE--
B. subtilis TKIYPGENVG RGGDDTLFAK IDGTVKFERF G-RDRKKVSV YPVAQ--
B. cereus TKIYPGVNVG RGGDDTLYAK VDGVVRFERL G-RDRKQVSV YPVAQEA
Consistency *8656*4*** 4*5*4**8*7 6**4*7*834 *084684979 3554100

(b)


. . . . . 10 . . . . . .
80α major capsid MEQTQKLKL NLQHFAS--
80α scaffold M-EENKLKF NLQFFADQS
S. aureus L27 -----MLKL NLQFFAS--
Consistency 30211 5**6 ***5**600

peptidyl transferase center and makes contact with the tRNA substrates (Selmer *et al.*, 2006). Sequencing of L27 from *T. thermophilus* by Edman degradation showed that the N-terminus of the peptide was alanine instead of formyl methionine (Kimura *et al.*, 1984). Furthermore, no density was detected for the N-terminal methionine of L27 in the high resolution structures of L27 (Voorhees *et al.*, 2009) suggesting it is not present.

Ribosomal protein L27 is highly conserved among different bacteria, chloroplasts of plants and red algae and the mitochondria of fungi. Sequence comparison of L27 from *E. coli* and *S. aureus* as well as several other Gram-positive and Gram-negative bacteria (Fig 25) reveals dissimilarity in the N-terminal region. Furthermore, peptide analysis of ribosomal proteins from a closely related Gram-positive bacterium, *B.subtilis*, performed using capillary LC−MS/MS, showed that the predicted N-terminal sequence, MLRLDLQFF, was missing in ribosomal protein L27 (Lauber *et al.*, 2009).

The absence of the N-terminal fragment suggests that this protein is also N-terminally cleaved between Phe and Ala. The sequences of ribosomal protein L27 from *S. aureus* and *B.subtilis* are highly similar and though no proteomic studies of staphylococcal L27 have been performed to date, it seems likely that staphylococcal L27 must be similarly cleaved by a staphylococcal protease. Furthermore, the similarity of the N-terminal region of ribosomal protein L27 and the scaffold and 80α major capsid proteins and the identical predicted P1-P1’ cleavage sites strongly suggest that these proteins are cleaved by the same host protease.

Discussion

Our studies clearly demonstrated that 80α scaffold and major capsid proteins are N-terminally cleaved and that the scissile bond lies between Phe and Ala within a

conserved sequence in both these proteins, indicating that a common protease is involved in proteolysis. Over-expression studies in *S. aureus* showed that 80α scaffold does not cleave major capsid protein or show any autocatalytic activity. Furthermore, a BLAST search with the scaffold protein sequence as a query did not show any similarity to known proteases or any conserved domain that could be attributed to any known family of peptidases.

Our observation that 80α and SaPI1 procapsids contained truncated scaffold and major capsid is in contrast to the accepted notion that cleavage triggers conformational changes leading to maturation of capsids as seen in HK97 (Gertsman *et al.*, 2009) and suggests that protein cleavage and procapsid maturation in 80α and SaPI1 particle assembly are connected events. The assembled procapsids represent the first irreversible stage of the procapsid assembly process and probably the role of N-terminal cleavage of scaffold and major capsid proteins is to prepare the procapsid for maturation events rather than be a cause of it. Most likely, the internal pressure generated by incoming DNA is the cause of procapsid expansion and maturation.

Inactivation of gp44 demonstrated that it does not have a proteolytic function, since no full length proteins were observed in the capsids obtained from mutant strains. A BLAST search using gp44 sequence as a query showed a distant similarity with phage SPP1 gp7, which is a neck protein. The neck proteins are found around the periphery of the portal complex which is incorporated at one of the vertices of the capsid. These proteins are present in minor amounts and play an important role in stabilizing capsids and preventing premature DNA ejection (Vinga *et al.*, 2006). Thus gp44 could be a neck protein, consistent with the observation that this protein is not

abundant in wt capsids. Growth studies with gp44 mutant strains showed that gp44 was essential for phage 80α even though virions were formed, implying that it confers stability to the 80α capsids and helps DNA ejection. SaPI1 particles obtained from the gp44 mutant strains transduced at wt levels and also seemed more stable, since they did not leak as much DNA as wt SaPI1 particles. The spatial arrangement of major capsid proteins around the portal complex is different in a T=7 capsid compared to a T=4 capsid. The role of gp44 in 80α capsids thus could be to stabilize the major capsid-portal protein interactions. The large number of free tails seen in the cryo-EM of ST64 would be consistent with this role. This reinforcement is not needed by SaPI1, owing to a different spatial arrangement in T=4 capsids. On the contrary, incorporation of gp44 in smaller capsids could be deterrent to the stability of smaller capsids and lead to DNA leakage.

The involvement of a host protease in N-terminal cleavage of 80α capsid and scaffolding proteins is novel and unexpected, adding a new mechanism to our knowledge of post-translational modification of capsid proteins in phages studied so far. In the absence of its own protease, 80α seems to use a host protease that is normally responsible for processing of ribosomal protein L27. L27 is an essential component of the 50S ribosome particle that is highly abundant in bacterial cells. Ribosomal protein L27 from Gram-negative bacteria, such as *E. coli*, have a formyl-methionine at the P1 position which is presumably removed by the action of two essential enzymes, peptide deformylase and methionine aminopeptidase (MAP) (Chang *et al.*, 1989; Arnold *et al.*, 1999; Giglione *et al.*, 2004). The loss of N-terminal methionine is the most common post translational modification found in bacteria through a proteolytic pathway known as N-

terminal methionine excision (NME) and these enzymes are probably responsible for removing f-Met from at least 34 other proteins (Arnold *et al.*, 1999) including ribosomal proteins (Ochi,1995). The N-terminus of the staphylococcal L27, however, has a prosequence of nine amino acids and it is unlikely that MAP has a role in its N-terminal processing. *S. aureus* appears to have a novel protease that is responsible for processing L27. 80α has devised a way to exploit this presumably essential host protease, thereby reducing its coding capacity and also ensuring high fidelity in terms of processing of its capsid and scaffold, by having a similar cleavage site and N-terminal sequence as the ribosomal protein L27.

S. aureus has a huge repertoire of proteases that consists of 131 known and putative peptidases and 43 non-peptidase homologues (Rawlings *et al.*, 2010) (http://merops.sanger.ac.uk/index.shtml). Serine peptidases account for almost one-third of all peptidases in *S. aureus*. The novel observation of a host protease being involved in the processing of major capsid and scaffold proteins raises many intriguing questions. First, what is the substrate specificity of this protease? Substrate specificity is attributed to the substrate binding pocket of an enzyme. Based on the consensus sequence KLKxNLQxF*A and identical P1-P1' sites in its substrates, this protease is an endopeptidase that recognizes extended substrates and/or a bulky aromatic residue at its S1 site. Very few enzymes are known that have extended specificity up to the P9 residue. Human caspase-3 and staphylococcal GluC are some proteases that have extended specificity. On the other hand, recognition of a bulky aromatic residue such as phenylalanine at the P1 position in the substrate is a feature commonly found in chymotrypsin-like serine endopeptidases. Do these N-terminally processed proteins

really need a protease that has stringent specificity when the problem of multiple cleavage can be easily averted by hiding other potential cleavage sites deep inside their structure? Second, what is the efficiency of this protease? Viral proteases do not have high catalytic efficiency (k_{cat}/K_m) and it has been argued that low k_{cat}/K_m value is beneficial for the self-assembly process (Babe *et al.*, 1997). On the other hand, bacterial proteases often have high catalytic efficiency. Interestingly, the major capsid and scaffold proteins have a different amino acid at the P2 position. Does this mean that the rate of cleavage of scaffold is different than of major capsid? We never detected either full length protein in our study with the procapsids, but that does not imply that the catalytic efficiency of the protease is high.

Our finding that host proteases could be involved in N-terminal processing of major capsid and scaffold adds a new dimension to the strategies that viruses employ to exploit their host. Furthermore, our study also strongly suggests that the staphylococcal ribosomal protein L27 is similarly processed, which is an important contribution to the understanding of the most fundamental and important mechanism of the translation process in *S. aureus*. Further studies are needed to identify this protease and study its substrate specificity. If this protease has a stringent specificity, then such studies could help design competitive inhibitors that target this host protease and thus provide a new weapon in our fight against *S. aureus*.

Chapter 5

Role of SaPI proteins in capsid size determination

Introduction

The production of smaller T=4 transducing particles is a striking characteristic of helper phage exploitation by SaPIs. In earlier studies in our lab (Tallent *et al.*, 2007), protein profiling of structural proteins of both sized particles by SDS-PAGE analysis failed to reveal any prominent bands that could be attributed to SaPI1-encoded proteins. Furthermore, in that study, intact phage particles were disrupted, digested with trypsin and peptide mass fingerprinting was done by mass spectrometry-MALDI-TOF. Mass spec analysis confirmed that the smaller sized particles were entirely composed of helper phage encoded proteins. This suggested that SaPI1 must act at an earlier step in capsid assembly to modify the capsid size. In order to investigate the possibility of a transient involvement of SaPI1 encoded proteins during their assembly, it was necessary to analyze the protein composition of both types of particles at an intermediate stage of their assembly rather than the final stage. In order to arrest self-assembly at the procapsid stage, small terminase subunits were deleted from 80α and SaPI1 (Chapter 3). Procapsid production was induced in 80α *ΔterS* prophage containing strain (ST24) and SaPI1 *ΔterS* positive 80α lysogenic strain(ST37) . Procapsids were purified on CsCl followed by sucrose gradients as described previously.

SDS-PAGE analysis of CsCl and sucrose gradient-purified procapsid proteins

SDS-PAGE analysis of CsCl-purified 80α and SaPI1 procapsids showed a set of bands that mostly corresponded to those that were identified in the earlier study by our lab (Tallent *et al.*, 2007). These bands included a full complement of head and tail proteins (Fig 26) consisting of minor tail proteins gp59, gp61, gp62 and gp68 (theoretical masses, 71.0, 73.7, 66.8 and 43.8 kDa, respectively), portal protein gp42 (59.5kDa), major capsid protein gp47 (36.8kDa) and major tail protein gp53 (21.5kDa). In addition, both 80α and SaPI1 procapsid samples had a prominent band at an apparent mass of approximately 27kDa that corresponds to the gp46 scaffolding protein (calculated mass, 23.4kDa). SDS-PAGE of the sucrose-gradient purified 80α and SaPI1 procapsids showed a similar set of capsid proteins, but greatly reduced amounts of the tail proteins. In addition, the SaPI1 procapsids yielded a faint band at 8kDa (Fig. 26), identified as SaPI1 gp6 by peptide mass fingerprinting.

Peptide mass fingerprinting of 80α and SapI1 procapsids

Peptide mass fingerprinting (PMF) (James *et al.*, 1993; Pappin *et al.*, 1993; Yates *et al.*, 1993) was employed for conclusive identification of structural proteins of 80α and SaPI1 procapsids. Proteins from CsCl-purified and sucrose gradient-purified procapsids were resolved by SDS-PAGE, in-gel digested with trypsin and analyzed by liquid chromatography and tandem mass spectroscopy (ESI-TOF). The PMF spectra of peptides were compared to an *S. aureus* protein database and all 80α and SaPI1 related proteins in the digest were identified. A comprehensive list of all 80α and SaPI1 proteins detected during the study is listed in Table 3

Figure 26. SDS-PAGE of the 80α and SaPI1

CsCl-purified protein fractions, as well as SaPI1 procapsids purified on sucrose gradients. The identities of major bands are indicated by the gene product number

. Reprinted from J Mol Biol. 380 (3) Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids, 465-475, (2008) with permission from Elsevier

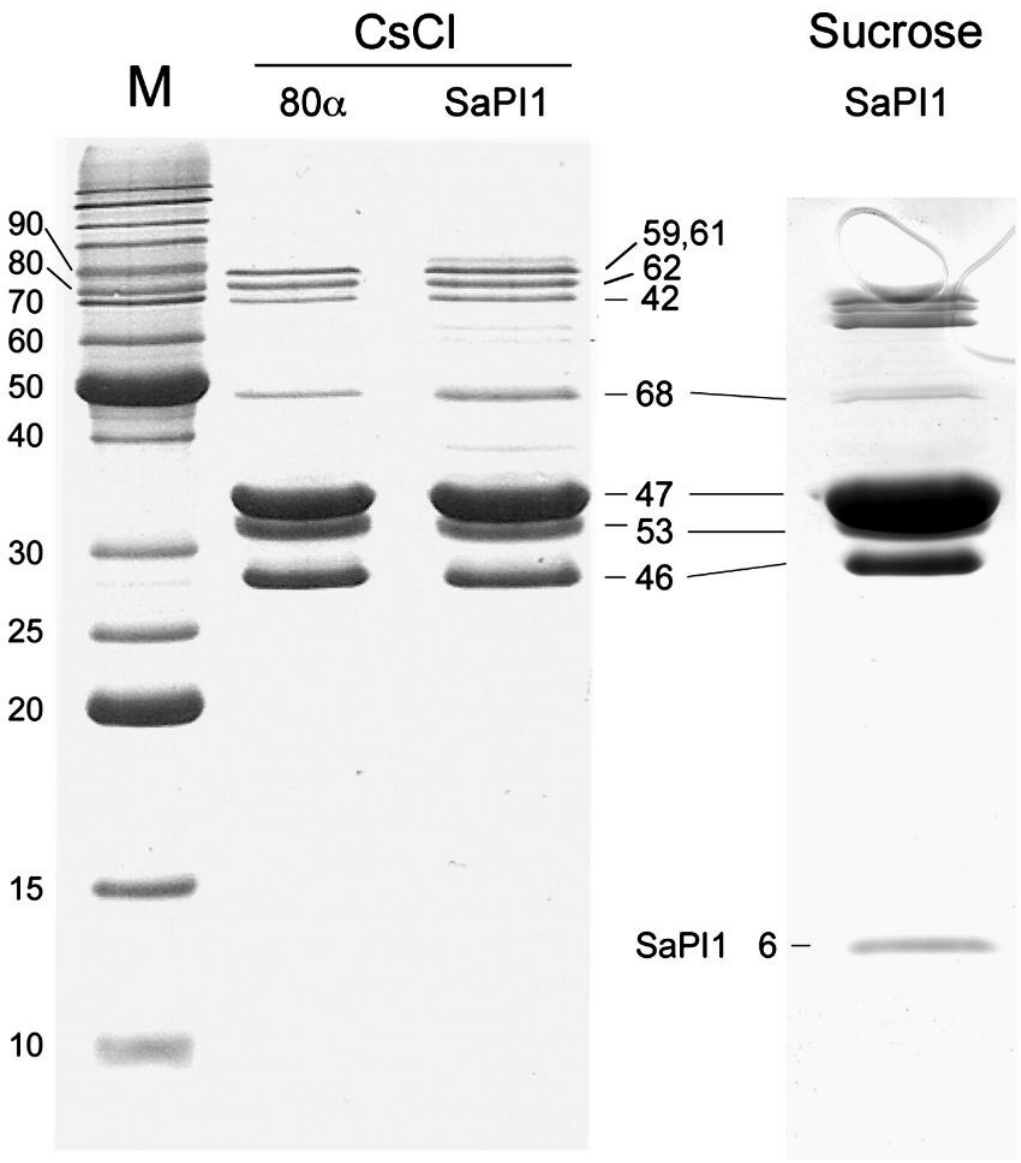

CsCl
Sucrose
M
80α
SaPI1
SaPI1
90
80
70
60
50
40
30
25
20
15
10
59,61
62
42
68
47
53
46
SaPI1 6

Table 3. Listing of all proteins detected by ESI-MS of trypsin digests

orf (gp)	Accession number	Number of amino acids	Molecular mass (kDa)	Function/location	% Sequence coverage			
					CsCl fraction		Procapsids	
					80α	SaPI1	80α	SaPI1
80α proteins								
03	YP_001285317.1	301	35.6	Unknown	6.0	20.6	–	–
08	YP_001285322.1	262	30.1	Putative antirepressor	40.1	39.7	–	–
10	YP_001285324.1	58	6.7	Unknown	–	22.4	–	29.3
15	YP_001285329.1	73	8.6	Unknown	19.2	–	–	–
16	YP_001285330.1	207	23.7	Unknown	27.5	30.9	5.3	9.7
17	YP_001285331.1	142	16.0	Single-stranded-DNA-binding protein	8.5	35.2	–	–
20	YP_001285334.1	256	29.7	Putative replisome organizer	12.5	–	–	–
32	YP_001285346.1	170	19.0	dUTPase	–	37.1	17.1	–
36	YP_001285350.1	128	15.0	Unknown	26.6	46.9	8.6	33.6
39	YP_001285353.1	140	16.5	Transcription activator (rinA)	–	9.3	–	–
40	YP_001285354.1	146	16.8	Small terminase subunit	–	29.5	–	–
41	YP_001285355.1	447	52.5	Large terminase subunit	–	13.0	–	–
42	YP_001285356.1	511	59.5	Portal protein	47.0	59.7	49.5	48.5
44	YP_001285358.1	331	38.5	Minor capsid protein, putative protease	37.1	38.1	30.0	28.2
46	YP_001285360.1	206	23.4	Scaffolding protein	47.6	74.8	67.0	54.4
47	YP_001285361.1	324	36.8	Major capsid protein	71.0	70.4	71.3	73.5
48	YP_001285362.1	95	10.9	Unknown	23.2	61.1	–	–

orf (gp)	Accession number	Number of amino acids	Molecular mass (kDa)	Function/location	% Sequence coverage			
					CsCl fraction		Procapsids	
					80α	SaPI1	80α	SaPI1
49	YP_001285363.1	110	12.8	Putative DNA-packaging protein	33.6	28.2	40.0	37.3
50	YP_001285364.1	100	11.8	Unknown	20.0	41.0	44.0	–
52	YP_001285366.1	127	14.7	Unknown	28.3	31.5	30.7	–
53	YP_001285367.1	193	21.5	Major tail protein	75.6	54.9	42.5	62.7
54	YP_001285368.1	121	13.1	Putative tail assembly protein	19.0	–	12.4	–
55	YP_001285369.1	114	13.6	Putative tail assembly protein	42.1	–	–	–
56	YP_001285370.1	1154	125.8	Tail tape measure protein	19.2	11.6	–	2.1
58	YP_001285372.1	315	37.1	Likely minor tail protein	61.3	51.7	5.7	35.2
59	YP_001285373.1	633	71.0	Lipase, likely tail tip	53.2	47.6	25.9	30.3
61	YP_001285375.1	636	73.7	Minor tail protein	24.7	67.8	28.9	54.9
62	YP_001285376.1	607	66.8	Minor tail protein, baseplate	37.9	67.2	35.0	56.2
64	YP_001285378.1	125	14.1	Likely minor tail protein	33.6	54.4	–	–
66	YP_001285380.1	99	11.8	Likely minor tail protein	–	17.3	–	–
67	YP_001285381.1	632	72.0	Cell wall hydrolase, likely tail tip	22.3	29.9	5.1	–
68	YP_001285382.1	390	43.8	Putative tail fiber protein	70.8	76.2	36.9	52.6
69	YP_001285383.1	131	14.4	Likely minor tail protein	67.9	58.8	42.7	60.3
71	YP_001285385.1	484	53.8	Lysis protein	13.6	–		–

orf (gp)	Accession number	Number of amino acids	Molecular mass (kDa)	Function/location	% Sequence coverage			
					CsCl fraction		Procapsids	
					80α	SaPI1	80α	SaPI1
SaPI1 proteins								
5	AAC28956.2	175	20.6	Putative accessory capsid protein	–	32.6	–	–
6	AAC28957.2	72	8.3	Putative capsid size determinant	–	91.7	–	91.7
7	AAC28958.2	192	22.8	Putative capsid size determinant	–	79.7	–	9.9
11	AAL67613.1	131	15.2	Unknown	–	16.8	–	–
18	AAL67617.1	106	12.6	Unknown	–	36.8	–	–

80α proteins identified in this study

All proteins previously identified in 80α and SaPI1 virions (Tallent *et al.*, 2007) were detected in this study. Additionally, minor amounts of some proteins not found in the earlier study were also identified. Several proteins were found in CsCl purified samples but not in sucrose purified samples or their presence was reduced in the samples from sucrose gradients. Most of these were components of tails which had co-purified with procapsids on CsCl gradients. Other proteins that were abundant in CsCl-purified samples but not in sucrose-purified procapsids corresponded to non-structural proteins. For example, gp8 is a putative antirepressor, gp71 is an endolysin and involved in cell lysis, gp20 is an initiator involved in DNA replication and gp48 has unknown function but since it belongs to the morphogenetic operon, it could have structural role. It is also quite possible that some of these nonstructural proteins have moonlighting functions. A number of host proteins were also detected in the CsCl-purified samples. Most of the host proteins that were identified were ones that are normally present in great abundance, for example ribosomal proteins, or enzymes with large molecular masses such as DNA polymerase. Additionally, we identified three more proteins in this study that were not previously identified (Tallent *et al.*, 2007). 80α protein gp10 (6.7kDa) of unknown function was found in SaPI1 procapsids, but not in 80α procapsids. Similarly, 80α gp16 (23.7 kDa) and gp36 (15.0 kDa) correspond to proteins of unknown function were found in both 80α and SaPI1 procapsids. The amount of these proteins in sucrose purified procapsids, except for gp36 in SaPI1 procapsids, was low suggesting that these proteins might be loosely associated with procapsids.

80α TerS (gp40) and TerL (gp41) were present in the SaPI1 CsCl sample, but not in sucrose-gradient purified SaPI1 procapsids. We also identified some tail proteins that were not previously identified. gp54 and gp55 were present in the 80α CsCl-purified protein sample and have putative tail assembly functions (Xu *et al.*, 2004). Also, gp58, gp64, gp66 and gp67, not previously detected, also most likely correspond to tail proteins.

SaPI1 proteins identified in this study

Five additional SaPI1-encoded proteins were found in the CsCl-purified SaPI1 protein sample that were not detected in SaPI virions. These are SaPI1 gp5, gp6, gp7, gp11 and gp18. Of all these proteins, only gp6, with an apparent mass of 8KDa, was found in abundance during SDS-PAGE analysis (Fig 26) of SaPI1 sucrose-purified procapsids, while gp7, found in CsCl gradient samples, was present in low amounts in sucrose purified samples. SaPI1 proteins gp5 and gp7 were not observed on the polyacrylamide gel during SDS-PAGE of the CsCl material because they co-migrate with the 80α scaffold, but their presence was confirmed by MALDI-TOF mass spectrometry. Of all the above mentioned proteins, the presence of gp6 and gp7 in procapsid samples was of most significance to this study. The homologs of these two proteins in SaPIbov1 had been previously implicated in size determination (Ubeda *et al.*, 2007; Ubeda *et al.*, 2008). Deletion of either of these genes eliminated the appearance of SaPI virion sized DNA, suggesting a defect in small capsid formation. The functions of gp5, gp11 and gp18 are unknown, but deletion of the homologous genes from SaPIbov1 did not affect SaPI replication or transduction.

Density measurements on the SDS-PAGE gels indicate that scaffolding and capsid proteins are found in a molar ratio of between 0.33:1 and 0.76:1, corresponding to 137–315 copies of gp46 and 415 copies of gp47, assuming T = 7 symmetry for the 80α capsid and one unique vertex. In SaPI1, the ratio is only 0.2:1, suggesting that other factors may substitute for some of the gp46. However, the SaPI1 measurements are complicated by the presence of expanded shells that had no scaffold.

Capsid size determination

The most important finding in these studies was identification of several SaPI1 proteins in the SaPI1 procapsids. In a previous study with SaPIbov1, proteins encoded by orfs 8 and 9 had been shown, with help of genetic experiments, to be involved in size determination (Ubeda *et al.*, 2007). SaPIbov1 gp8 and gp9 are homologs of SaPI1 gp6 and gp7, respectively. Both gp6 and gp7 were abundant in the CsCl gradient purified sample and detected by ESI-TOF mass spectrometry, however only gp6 was present in abundance in sucrose purified procapsids suggesting gp7 is loosely attached to procapsids. The presence of gp6 and gp7 in procapsids but not in mature SaPI1 capsids suggests that they are transiently associated with procapsids and most likely function as alternate scaffolding proteins in capsid size redirection.

Comparing SaPI1 size determination proteins with P4 encoded Sid protein

These experiments clearly established that SaPI1 sized particles are formed because of a transient involvement of SaPI1 encoded proteins. 3D-reconstructions of cryo-EM further established that the icosahedral symmetry of the SaPI1 particles is T=4 while that of 80α particles is T=7. The predicted secondary structures of both gp6 and

gp7 are α-helical which is characteristic of scaffolds. Thus, there are striking similarities between these two systems. So are SaPI1 gp6 and gp7 external scaffolds that use the same mechanism as Sid to subvert the T=7 pathway? The cryo-EM studies showed that the external surface of SaPI1 procapsids, unlike the surface of P4 procapsids, is smooth and similar to that of 80α procapsids. In P4 procapsids, Sid is clearly visible as angular density on the procapsid exterior. Thus these data suggest that SaPI1 uses an internal scaffold to redirect capsid assembly.

We took the following approaches to investigate if SaPI1, gp6 and gp7 are sufficient and necessary for capsid size redirection. First, in order to confirm that SaPI1 gp6 and gp7 are necessary for capsid size redirection, as was the case in SaPIbov1, we performed loss-of-function studies in SaPI1 by deleting orf6 and orf7 and observing their phenotypes by electron microscopy and agarose gel electrophoresis of encapsidated DNA. Furthermore, we also wanted to ascertain whether these proteins are sufficient for capsid size redirection and rule out the possibility of involvement of other SaPI1 proteins by expressing these two proteins alone during an 80α infection. In order to ensure that the SaPI1 proteins are cotemporally expressed and have comparable stoichiometry to 80α head proteins, we constructed transcriptional fusions by inserting SaPI1 orfs 6 and 7 into the 80α morphogenetic cluster downstream of 80α orf44. The native ribosome binding sites of these two genes were retained to ensure correct stoichiometry of the expressed proteins.

Results

SaPI1 orf6 and orf7 deletion mutants were generated by allelic exchange with plasmid pMAD derivatives pPD35 and pPD34, containing flanking regions of their

respective genes, in SaPI1-positive 80α lysogenic strain RN10628, generating strains ST100 and ST98, respectively. Transcriptional fusions were constructed by inserting SaPI1 orf6 and orf7 individually and in conjunction, by allelic exchange, using pMAD derivatives pPD36, pPD37 and pPD28 respectively. These were inserted between 80α orfs 44 and 45 in the morphogenetic operon at nucleotide position 20694 of an 80α prophage in lysogenic strain RN10616 to generate strains ST97, ST99 and ST82, respectively. All of these strains were induced by mitomycin C and their lysates were examined for capsid morphology by cryo-electron microscopy and for capsid volume by agarose gel electrophoresis of the encapsidated DNA. In addition, plaque and transduction assays (Novick,1991; Kropinski *et al.*, 2008) of the resultant lysates were performed along with their appropriate controls to investigate the role of capsid size redirection in interference with 80α propagation.

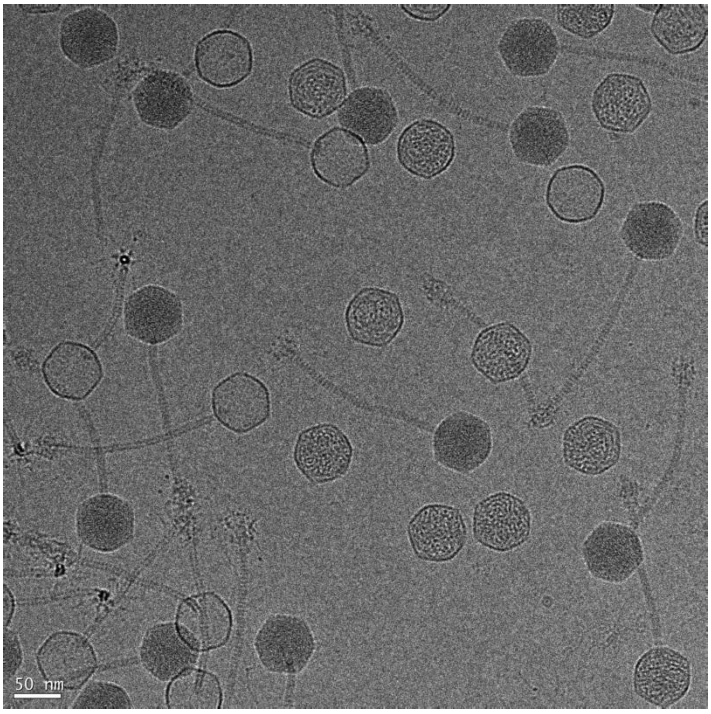
SaPI1 gp6 and gp7 are both necessary for formation of small capsids

Examination by cryo-EM of CsCl banded particles generated by induction of a strain carrying the SaPI1 Δorf7 mutant (Fig 27a) revealed a mixture of filled and empty particles that were all large. These particles had isometric capsids similar to expanded mature 80α capsids in size. About 22% of the shells were empty and presumably had lost their DNA during or after purification. Typically, in a lysate of a wt SaPI1-positive 80α lysogen, about 98% of the particles are small. In contrast, small particles were completely absent in these mutant SaPI1 samples suggesting that capsid size redirection was completely abolished by deletion of SaPI1 orf7. Cryo-EM analysis of CsCl-banded particles formed by the SaPI1 Δorf6 mutant (Fig 27b) yielded intriguing

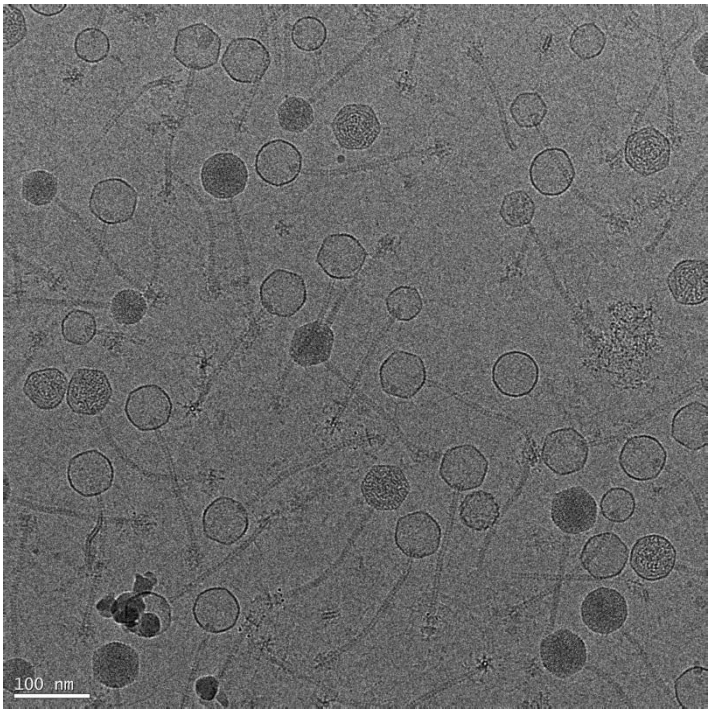
Figure 27. Effects of deleting SaPI1 orf6 or SaPI1 orf7 .

Lysates from SaPI1 Δorf7 80α lysogen (ST98) (a) or SaPI1 Δorf6 80α lysogen (ST100) (b) were banded in CsCl prior to visualization by cryo-EM.

(a) SaPI1 Δorf7



(b)SaPI1 Δorf6



results. There were empty (56%) and full (44%) particles having isometric capsids of large and small sizes. Only 51% of small capsids (25%) and about 40% of the large capsids (62%) were full. Furthermore, there were about 13% aberrant capsids that did not conform to T=4 or T=7 symmetries. These aberrant capsids were mostly prolate although some oblate capsids were also observed. Thus deletion of SaPI1 orf6 affected capsid size redirection and considerably reduced the number of small capsids. Furthermore, fidelity of small capsids was also affected as particles with aberrant morphologies not previously seen were observed in these samples.

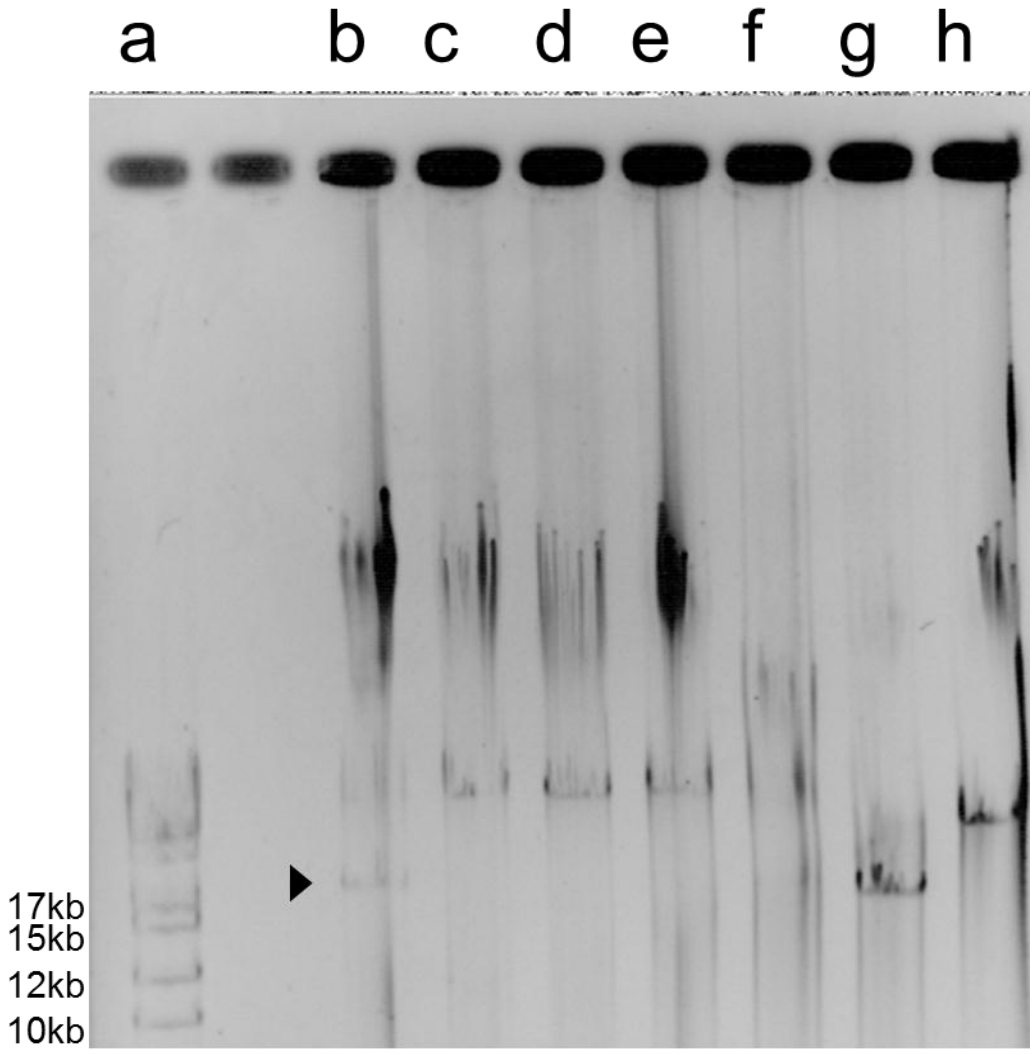
Agarose gel electrophoresis of the DNA from the minilysates for both these deletion mutants (Fig 28) supported the above observations. Both these mutants showed a loss of the faster migrating fragment implying that small capsids were not formed. Thus, from the above data it is inferred that both SaPI1 gp6 and gp7 are necessary for capsid size redirection and that deletion of either orf results in loss of small capsid formation.

SaPI1 gp6 and gp7 are sufficient for capsid size redirection

In order to ascertain if SaPI1 gp6 and gp7 are sufficient for small capsid formation, we analyzed lysates from strains ST97, ST99 and ST82, which had transcriptional fusions of orf6, orf7 and orf6 and 7 inserted into the 80α capsid gene cluster. CsCl-banded particles in lysates resulting from induction of these strains were examined by cryo-electron microscopy and agarose gel electrophoresis.

Figure 28. Agarose gel electrophoresis of encapsidated DNA from SaPI1 deletion and insertion strains

0.7% Agarose gel showing migration of monomer sized DNA obtained from lysates of the following insertion and deletion mutants.
(a) molecular ladder
(b) 80α::orf 6+7
(c) 80α::orf 7
(d) 80α::orf 6
(e) 80α + SaPI1Δorf7
(f) 80α + SaPI1Δorf6
(g) 80α + SaPI1
(h) 80α
The SaPI1 monomer sized band is indicated with an arrow


a
b
c
d
e
f
g
h
17kb
15kb
12kb
10kb

Cryo-EM analysis of particles generated by the orf6 insertion mutant (Fig 29a) revealed a mixture of particles that were predominantly large, with only about 3% capsids that were small. Furthermore, about 80% of the particles were filled with DNA indicating that the particles were stable.

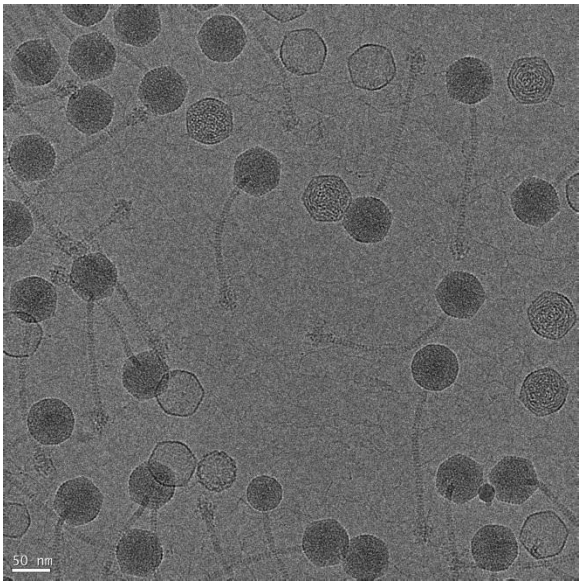
Cryo-EM analysis of CsCl-banded particles produced by the orf 7 insertion mutant (Fig 29b) revealed particles with only large capsids. Small capsids were completely absent in these samples. However, nearly half of the large capsids were empty. As with the orf6 insertion, agarose gel electrophoresis of the encapsidated DNA from the orf7 insertion (Fig 28) also did not show smaller SaPI1 sized bands. Thus, SaPI1 gp6 or gp7 alone were not sufficient for formation of smaller sized SaPI capsids.

Since SaPI1 gp6 and gp7 were both required for capsid size redirection we also investigated whether both proteins, in conjunction, are sufficient for formation of small sized capsids. Cryo-EM analysis of CsCl-banded particles resulting from strains carrying an insertion of orfs 6 and 7 in conjunction (Fig 29c) revealed that the particles were predominantly of smaller size with T=4 symmetry; only 2% were large. Furthermore, agarose gel electrophoresis of the encapsidated DNA (Fig 28) also showed the presence of the smaller SaPI sized DNA fragment implying that small capsids were being formed. Thus SaPI1 gp6 and gp7 together are sufficient to redirect capsid size to form smaller sized capsids. The above experiments allow us to conclude that both SaPI1 gp6 and gp7 are necessary and sufficient for forming small capsids.

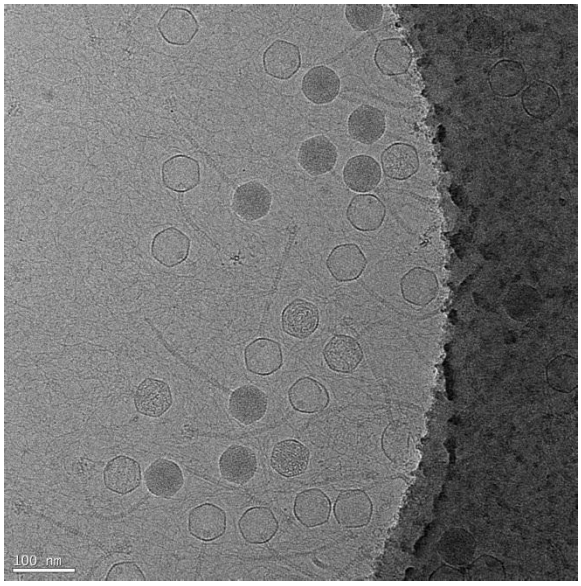
Figure 29. EM analysis of 80α mutants carrying insertions of SaPI1 gp6 and/or gp7.

Cryo-electron micrographs of CsCl purified particles obtained from lysogenic 80α strains that had SaPI1 orf6 (A) orf7 (B) and orf6 and orf7 together (C) inserted in them.

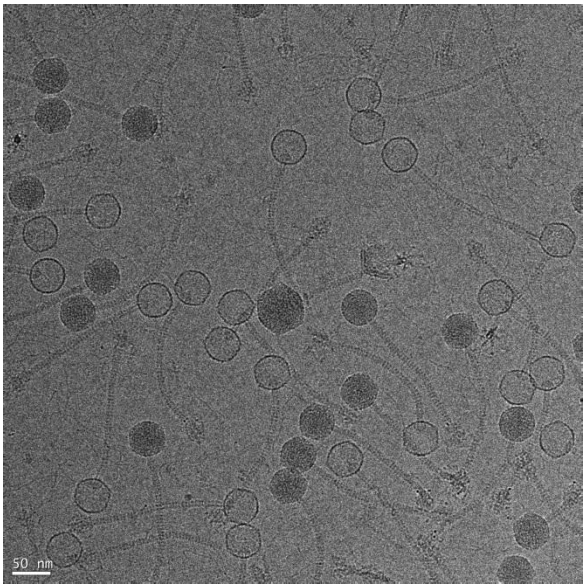
A 80α ::SaPI1 orf6



B 80α ::SaPI1 orf7



C 80α ::SaPI1 orf 6+7



How does capsid size redirection affect 80α propagation?

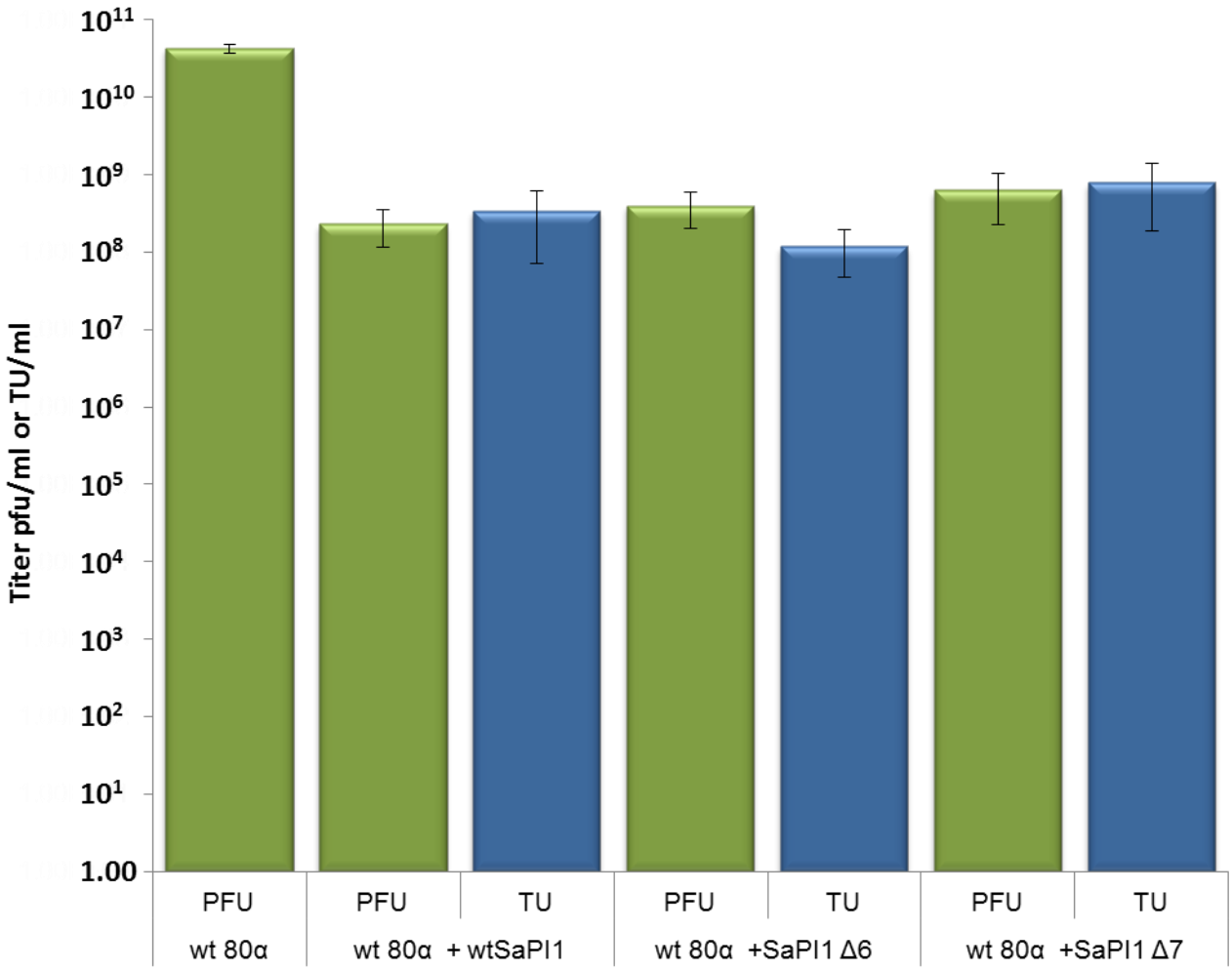
Another notable characteristic in the 80α–SaPI1 relationship is interference with the propagation of 80α in the presence of SaPI1. Induction of a SaPI1 positive 80α lysogen yields a lysate having a phage titer that is about two orders of magnitude lower than a lysate of 80α alone. This reduction in burst size is sufficient to block 80α plaque formation on a lawn of SaPI1-positive *S. aureus*. Similar interference in the P2/P4 system has been attributed to capsid size redirection (Diana *et al.*, 1978; Nilssen *et al.*, 1996). In SaPI mobilization, there are several possible factors contributing to helper phage interference including altered packaging specificity due to the SaPI1 encoded small terminase and formation of small capsids that physically limit encapsidation of a full length 80α genome. Earlier studies in our lab demonstrated that deletion of SaPI1 *terS* decreased SaPI1 transduction by at least three orders of magnitude but did not relieve interference (Ubeda *et al.*, 2009). We expected that small capsid formation would interfere with 80α propagation due to the inability to package a complete 80α genome. Therefore one of the goals of this study was to find out the correlation between capsid size redirection and interference in 80α propagation. We performed growth studies with these mutants to determine if blocking small capsid formation would relieve the phenomenon of interference or whether small capsid formation per se caused interference.

Results**:**

Phage and transduction assays were performed according to methods outlined in

Figure 30. Effects of SaPI1 deletion mutants on phage and SaPI1 titers

Plaque assays of 80α and transducing titers of SaPI1 of lysates resulting from SaPI1 orf6 (ST100) and orf7 (ST98) deletion mutant strains. The mutant strains ST98 and ST100 were induced with Mitomycin C and allowed to lyse. Phage and SaPI transduction titers were assayed on a sensitive *S. aureus* indicator strain,RN4220.
The plaque titers are shown in green and the transducing titers are in blue. Three biological replicates were used in this assay.


Titer pfu/ml or TU/ml
10^11
10^10
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
1.00
PFU
wt 80α
PFU
TU
wt 80α + wtSaPI1
PFU
TU
wt 80α +SaPI1 Δ6
PFU
TU
wt 80α +SaPI1 Δ7

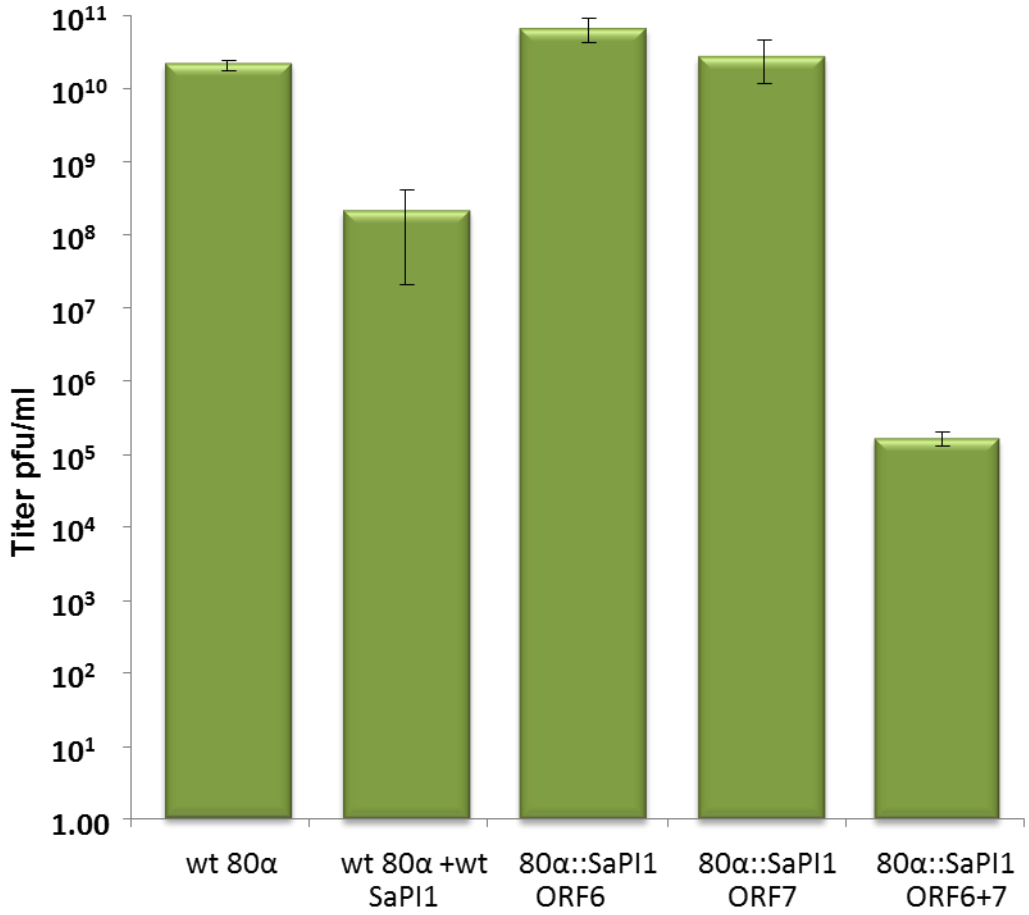
Materials and Methods. Although deletion of SaPI1 orfs 6 or 7 did not alter capsid size, we expected them to relieve interference. Surprisingly, however, they did not. The phage titers from these deletion strains were still significantly lower than the wt 80α indicating that interference was not relieved (Fig 30). Deleting orf6 or 7 did not affect SaPI1 transduction frequencies either (Fig 30). The SaPI1 transduction frequency was comparable to wild type, suggesting that SaPI1 DNA was efficiently encapsidated in large capsids in the absence of small ones and that SaPI1 did not preferentially package its DNA in small capsids.

We also performed plaque assays with the insertion strains (Fig 31). 80α carrying an insertion of either SaPI1 orf6 or orf7 alone formed plaques at wild type levels. The phage titer obtained by induction of a lysogen carrying an insertion of both of these genes dropped sharply to five orders of magnitude less than the wild type. We expected the phage titer of this mutant strain to be below detectable limits, similar to the plating of 80α on SaPI1 positive strains. The relatively high residual phage titer could be explained by a high number of spontaneous inactivation mutations occurring in these two genes due to repeated selection during each burst (multiplication) cycle. Subsequent analysis of some of these plaques is consistent with this interpretation.

Coexpression of SaPI1 genes 6 and 7 with 80a genes resulted in capsid size redirection and also interfered with 80α growth, consistent with our hypothesis. However, deletion of either SaPI1 gene 6 or 7 blocked small capsid formation, but did not relieve interference with helper phage growth. This result was surprising and it

Figure 31. Effects of inserting SaPI1 orf6 and/or orf7 on phage titers

Plaque assays of 80α carrying SaPI1 orf 6 and/or 7 insertion. The 80α lysogenic strains containing insertions of SaPI1 orf6 and 7(ST82), SaPI1 orf7 (ST97) and SaPI1 orf6 (ST99) were induced with Mitomycin C and allowed to lyse
Phage titers were assayed on a sensitive *S. aureus* indicator strain,RN4220.


Titer pfu/ml
10^11
10^10
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
1.00
wt 80α
wt 80α +wt SaPI1
80α::SaPI1 ORF6
80α::SaPI1 ORF7
80α::SaPI1 ORF6+7

implies that in the absence of capsid size redirection, there are other factors that cause interference. DNA packaging specificity due to SaPI1 TerS is one such factor and there could be others as well. Thus the interference with helper phage 80α growth is a multifactorial event and other genes apart from SaPI1 encoded *terS* and capsid size determination genes are involved. Further studies aimed at identifying these genes and understanding their mechanism of involvement in interference would be needed to explain this effect.

This study clearly establishes the role of SaPI1 gp6 and gp7 proteins in capsid size redirection and shows that both are required along with 80α head proteins for efficient assembly of smaller sized particles. Capsid size redirection is not necessary for high frequency transduction of SaPI1 as SaPI1 can be efficiently packaged in large capsids as well. It is intriguing why both SaPI1 6 and 7 genes are highly conserved among staphylococcal pathogenicity islands. Deletion of these genes in SaPI1 did not affect SaPI1 transduction. Perhaps the answer to this puzzle lies in events that take place outside of the cell after lysis.

80α upon release after cell lysis infects its neighboring cells. Like all other temperate phages, it undergoes lytic life cycle, most of the times, instead of preferring to integrate into the S. aureus chromosome. This results in several subsequent life cycles, in the neighboring cells, causing an exponential increase of its progeny and a drastic reduction of the neighboring *S. aureus* population.

SaPI1 on the other hand lacks the propagation machinery and integrates into the *S. aureus* chromosome upon infection. Thus, in the event of release of phage and SaPI from an infected cell and subsequent infection of neighboring cells, the SaPI1

population could easily be outnumbered owing to the fact that phages are primed for multiplication while SaPIs are not. This could create a situation where the SaPI is completely lost and thus would be disadvantageous to SaPI1. SaPI1, by capsid size redirection, compels 80a to package incomplete genomes into the smaller particles. This molecular “castration” not only ensures that the number of viable phage particles is kept in check but also that SaPI1 has enough S. aureus population for its transduction.

Do SaPI1 genes 6 and 7 supplement or replace scaffold?

Having established the roles of SaPI proteins gp6 and gp7 in capsid size redirection, we also wanted to analyze if these proteins replaced or supplemented the function of 80α encoded scaffold. We generated gp46 mutant strains by deleting the scaffold gene in an 80α prophage, in SaPI1 positive and SaPI1 negative *S. aureus* strains, to generate ST51 and ST91, respectively. Growth studies were carried out on these strains. Furthermore, particles resulting from these strains were also subjected to EM analysis.

Results

80α encoded scaffold is essential for both 80α and SaPI1 particles

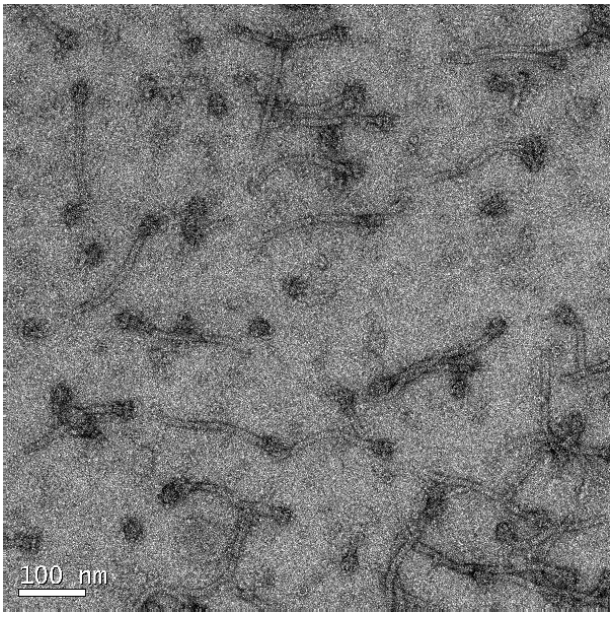
SaPI1 positive and negative gp46 inactivation mutant strains ST51 and ST91 were induced with mitomycin C and monitored until lysis. The particles obtained from these lysates were subjected to EM examination (Fig 32a). EM analysis of structures formed in a SaPI1 negative gp46 mutant strain failed to reveal any capsid-like structures of any morphology. The lysate from a SaPI1 positive gp46 mutant strain yielded some smaller sized procapsids and aberrant structures, in very low quantity. Furthermore,

there was an abundance of free tails and portal complexes in both these samples. These lysates were also assayed for viable phage and SaPI transducing particles. Both of these mutant strains failed to produce plaque forming units above detectable limits. Furthermore, the lysate resulting from the SaPI1-positive gp46 mutant did not yield any tetracycline resistant colonies. Thus, deletion of 80α orf46 had a drastic effect on the self-assembly process of both small and large capsids. No capsids were assembled at all in the absence of gp46 and SaPI1 (Fig 32a) suggesting that gp46 is very important for the nucleation event. Furthermore, only a very few small empty procapsids and aberrant structures assembled in the presence of SaPI1 (Fig 32b), indicating that gp46 is not only important for assembly of large but also small capsids. There were a lot of free tails and structures resembling portal rings seen in these EMs, indicating that gp46 is also needed for incorporating the portal proteins into the procapsid. From these data it can be deduced that the major capsid protein does not self-assemble even into aberrant structures, but requires the help of a scaffold for its assembly. SaPI1, proteins seem to allow some assembly of gp47 in the absence of gp46, but the yield of the particles is very low indicating that these two proteins supplement the scaffolding functions of gp46 rather than complement it. Furthermore, the 80α scaffolding protein also functions in incorporation of other minor proteins such as the portal proteins, as supported by the observation of free portal rings and free tails and also by our growth studies wherein we observed that deletion of orf46 was lethal for 80α as well as SaPI1.

Figure 32. Effects of deleting 80α scaffold on phage and SaPI1 particles

Electron micrograph of CsCl purified particles obtained from 80α Δorf46 lysogens induced in (a)SaP1 negative (ST91) and (b)SaPI1 positive (ST51).
Samples for negative stain were prepared by applying phage suspension to glow-discharged carbon-only grids (Electron Microscopy Sciences), washed 2× with dialysis buffer and stained with 1% uranyl acetate. Samples were observed in an FEI Tecnai F20 electron microscope operated at 200kV and images were captured on a 4k × 4k Gatan Ultrascan CCD camera or on Kodak SO-163 film at magnifications from 38,000× to 81,200×.

A



B

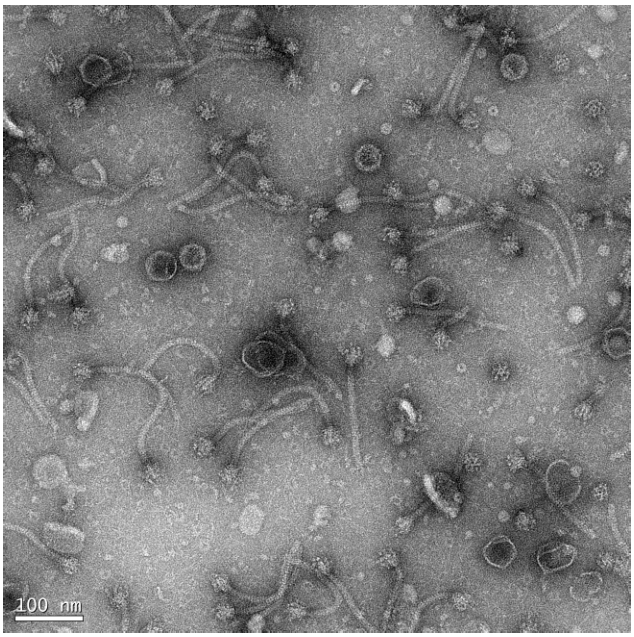
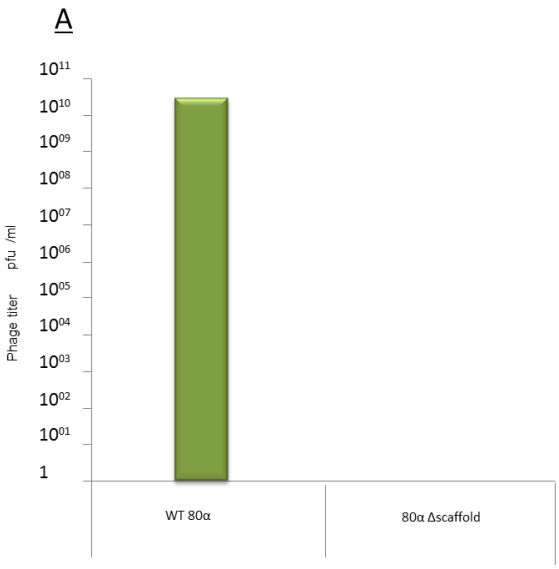
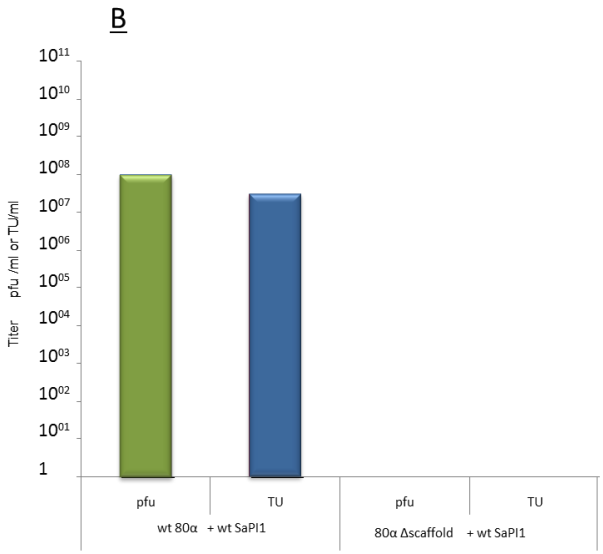


Figure 33. Effects on 80α Δorf46 phage and transducing titers.

Plaque and transduction assays of (a)SaP1 negative (ST91) and (b)SaPI1 positive (ST51) 80α*Δorf46* lysogenic strains. These strains were induced with Mitomycin C and allowed to lyse. Phage titers were assayed on a sensitive *S. aureus* indicator strain, RN4220. The plaque titers are shown in green, and transducing titers are in blue.


A
Phage titer pfu /ml
10^11
10^10
10^09
10^08
10^07
10^06
10^05
10^04
10^03
10^02
10^01
1
WT 80α
80α Δscaffold


B
Titer pfu /ml or TU/ml
10^11
10^10
10^09
10^08
10^07
10^06
10^05
10^04
10^03
10^02
10^01
1
pfu
TU
pfu
TU
wt 80α + wt SaPI1
80α Δscaffold + wt SaPI1

Chapter 6

Defining the minimum number of components required for capsid assembly

Introduction

The studies described thus far identified two SaPI1 proteins, namely SaPI1 gp6 and gp7, that were transiently associated in procapsids but not found in mature virions (Tallent *et al.*, 2007; Poliakov *et al.*, 2008). Furthermore, our genetic and EM studies with SaPI1 gp6 and gp7 established their role in size redirection (Chapter 6). We wished to further dissect the roles of the different proteins in capsid assembly by examining the process with a smaller number of components. In the P2/P4 system, in vivo expression studies showed that the P2 capsid protein gpN could assemble by itself into procapsid like particles with a triangulation number T=7, though with a low fidelity. Incorporation of the P2 scaffold gpO with gpN resulted in increased fidelity of the T=7 procapsid like particles. Addition of gpSid to this system now redirected capsid assembly, leading to formation of predominantly smaller procapsid like structures with a T=4 symmetry (Wang *et al.*, 2000). The effect of gpSid on capsid assembly was powerful enough to override the effect of the internal scaffold and redirect procapsid assembly to form smaller capsids. This chapter describes a set of experiments in the 80α-SaPI1 system.

Rationale

Virus assembly can be studied both in vitro and in vivo and both approaches have their own advantages and shortcomings. The in vitro approach relies on

expressing and purifying individual components, which are then added to the assembly reaction and subjected to thermodynamic and kinetic studies. One serious drawback of this approach is that it is difficult to establish an intracellular environment in a test tube and for this reason it might be easier to use the in vivo approach to determine the minimum subset. Given the success of this approach in the P2/P4 system, we undertook *in vivo* co-expression to determine the minimum subset of structural components of an 80α shell. Having determined the minimum components required for assembly of *bona fide* procapsid-like particles, we also wanted to determine how the SaPI1 encoded internal scaffolding proteins would interact with this subset and affect capsid size.

Our initial attempts to perform expression studies with 80α gp46 and gp47 using the T7 expression system in *E. coli* generated amorphous monster-like structures that were more than 100 nm in size, or sheet-like assemblies, instead of particles with a discrete morphology (Fig 4.2c). We initially narrowed this failure to produce procapsid-like structures to two main possibilities. (a) We observed that gp47 was insoluble, suggesting that most probably these proteins were misfolded in *E. coli,* and (b) there was a loss of N-terminal cleavage of both the scaffold and major capsid proteins when expressed in *E. coli* (Chapter 4). In order to address the problem of loss of proteolytic cleavage, truncated versions of major capsid and scaffold genes were cloned in tandem and expressed in *E. coli* in our collaborator's lab. However, the truncated proteins also failed to yield discrete procapsid-like particles, but still formed the amorphous structures previously observed with the full length proteins. Thus, the failure to form discrete procapsid-like particles was attributed to protein misfolding in *E. coli.* In order to

overcome this problem, we turned to a recently developed staphylococcal T7 expression system (D'Elia *et al.*, 2006; Pereira *et al.*, 2008) to express 80α proteins. 80α and SaPI1 proteins were expected to fold correctly in their natural host. We cloned 80α structural genes in various combinations onto pG164, a shuttle plasmid vector for the staphylococcal T7 expression system, and supplemented them with SaPI1 genes *6* and *7*. Two hours after induction with IPTG, the proteins expressed in cells carrying these plasmids were subjected to SDS-PAGE analysis and the lysates were subjected to cryo-EM analysis in Dr. Dokland's lab.

Results

Major capsid protein and scaffolding protein are both required for the self-assembly of procapsids

The major capsid protein gp47, on its own, did not assemble into procapsids when expressed in *S. aureus* strain ST70 (Fig. 34a). In addition, its expression level as visualized during SDS-PAGE analysis was low (Fig. 35). However, on co-expressing scaffolding protein with the major capsid, we observed some self-assembly into procapsids. EM analysis of the samples resulting from co-expression of scaffold with major capsid protein (Fig. 34b) revealed that about one-third of the total procapsid like structures resembled *bonafide* large (T=7) procapsids, about a third were expanded and the rest were aberrant assemblies with some super monster like structures. SDS-PAGE analysis showed that gp46 did not express well in this strain which perhaps explains the small number of T=7 procapsids and large number of aberrant structures. Co-expression of proteins from the plasmid containing the entire 80α orf42 to orf47 gene cluster generated procapsids with much higher fidelity (Fig 34c), suggesting that capsid

assembly was aided by additional proteins. These data suggest that the minimum subset required for self-assembly of 80α procapsids consists of scaffold and major capsid proteins, and that other proteins encoded by the head gene cluster assist in this process but are not absolutely essential.

SaPI1 gp6 and gp7 redirect capsid assembly

Having established the minimum subset required for assembling large (T=7) procapsids, we added SaPI1 genes 6 and 7 in tandem to the plasmid carrying the 80α scaffold and major capsid genes. Expression from this plasmid was induced with IPTG and the cells again analyzed after an induction period of 2 hours. SDS-PAGE analysis showed bands for gp47, gp46 and gp6 (Fig. 36). gp7 co-migrates with scaffold and hence could not be differentiated on the gel. EM analysis showed that small (T=4) procapsid like structures were now being assembled along with large (T=7) procapsids (Fig 37). There were a lot of aberrant structures as well. These data suggest that addition of SaPI1 gp6 and gp7 to the minimum subset of 80α proteins did result in capsid size redirection to form smaller procapsid like structures and that the minimum subset for formation of smaller procapsid consists of 80α scaffold and major capsid proteins supplemented with SaPI1 gp6 and gp7.

Discussion

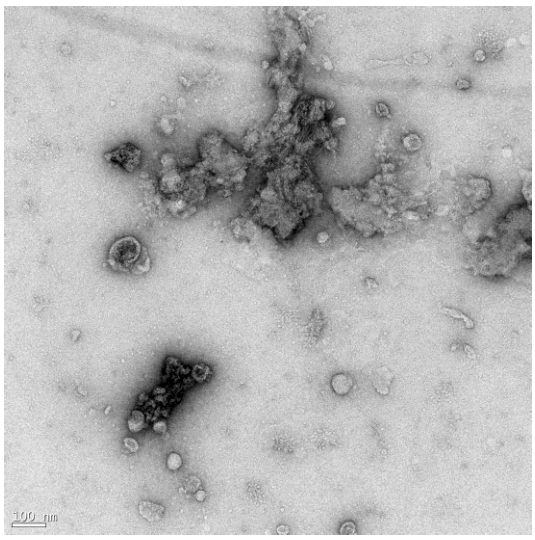
This study strongly suggests that 80α scaffold is very important for assembling large and small capsids and that 80α major capsid does not assemble correctly on its own. Studies performed with P22 have also shown that scaffold is very important for assembling particles of the correct size with high fidelity. Capsid assembly in the

Figure 34. Structures formed by coexpression of cloned 80α capsid genes.

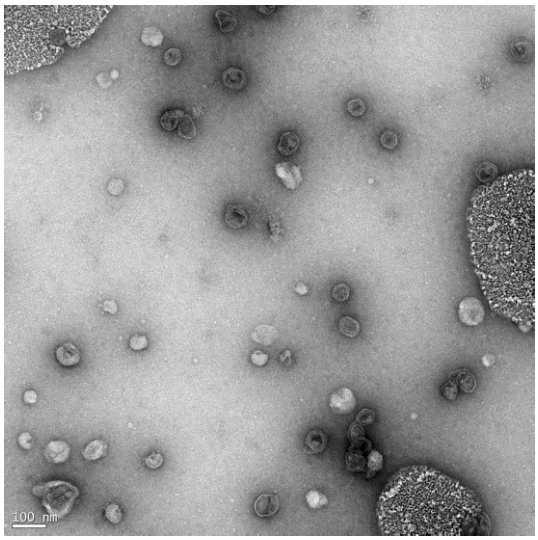
Major capsid protein was expressed from (A) ST70, scaffold and major capsid protein from (B) ST71 and portal through major capsid in (C) ST72.
Samples for negative stain were prepared by applying phage suspension to glow-discharged carbon-only grids (Electron Microscopy Sciences), washed 2× with dialysis buffer and stained with 1% uranyl acetate. Samples were observed in an FEI Tecnai F20 electron microscope operated at 200kV and images were captured on a 4k × 4k Gatan Ultrascan CCD camera or on Kodak SO-163 film at magnifications from 38,000× to 81,200×.

.

A



B



.

C

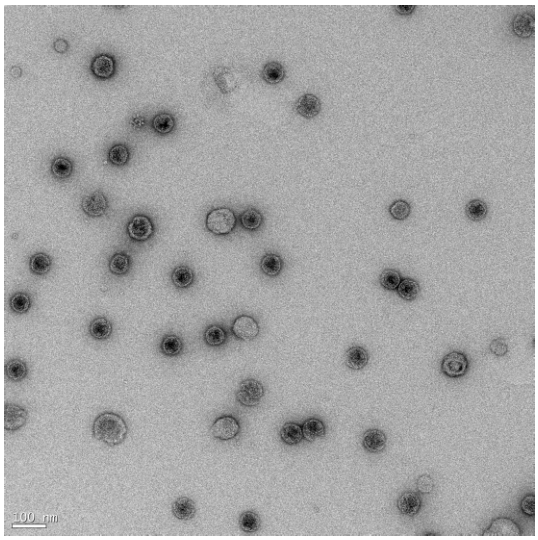


Figure 35.SDS-PAGE analysis of proteins obtained by coexpressing cloned 80α genes

Major capsid protein was expressed in ST66 from a plasmid with a native rbs and ST70 from a plasmid with engineered rbs, scaffold and major capsid protein from ST71 and portal through major capsid in ST72.
Cells were disrupted by boiling in reducing buffer for 10 minutes and loaded on 12% Bis-Tris precast gels (Criterion XT Biorad laboratories)

Lane (a) molecular weight marker size in KDa indicated
(b) wt 80α virions
(c) Plasmid pPD2, expressing major capsid and scaffold in *E. coli.*
(d) Plasmid pPD18, expressing major capsid from its native rbs in *S. aureus*
(e) Plasmid pPD20, expressing major capsid in *S. aureus*
(f) Plasmid pPD21 expressing major capsid and scaffold in *S. aureus*
(g) Plasmid pPD22 expressing portal through major capsid in *S. aureus*
(h) Plasmid pG164 as a negative control

Arrowheads indicate the bands corresponding to portal (in blue), capsid protein (in black) and scaffold protein (in red).

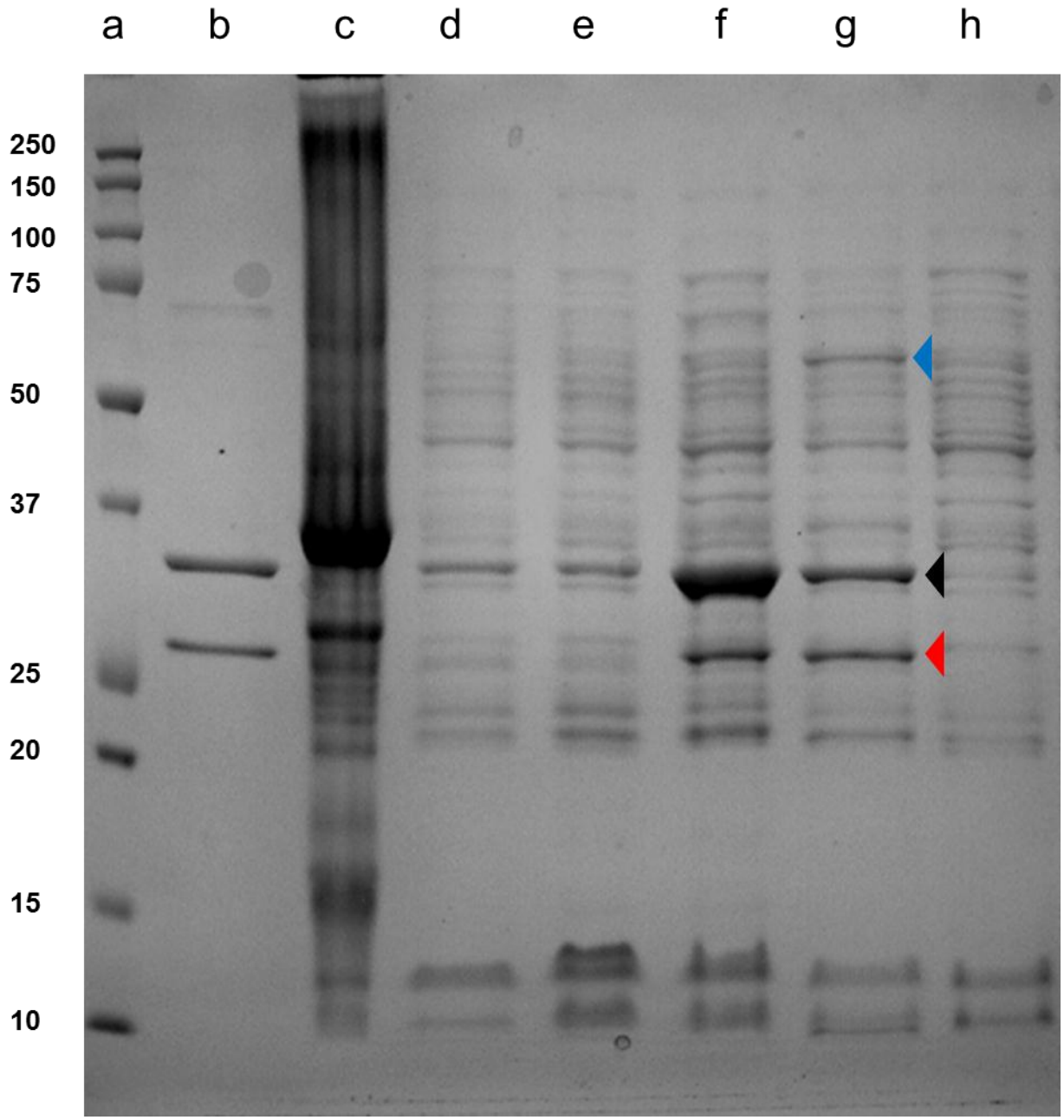

a
b
c
d
e
f
g
h
250
150
100
75
50
37
25
20
15
10

Figure 36. SDS-PAGE analysis of proteins obtained from coexpression of major capsid and scaffold along with SaPI1 6 and 7 genes.

Coexpression of cloned major capsid, scaffold SaPI1 genes 6 and 7(ST118) was done by induction by IPTG for 2 hours. Cells were lysed and samples were prepared for SDS-PAGE analysis
10% Tris HCl gel was run under denaturing conditions.

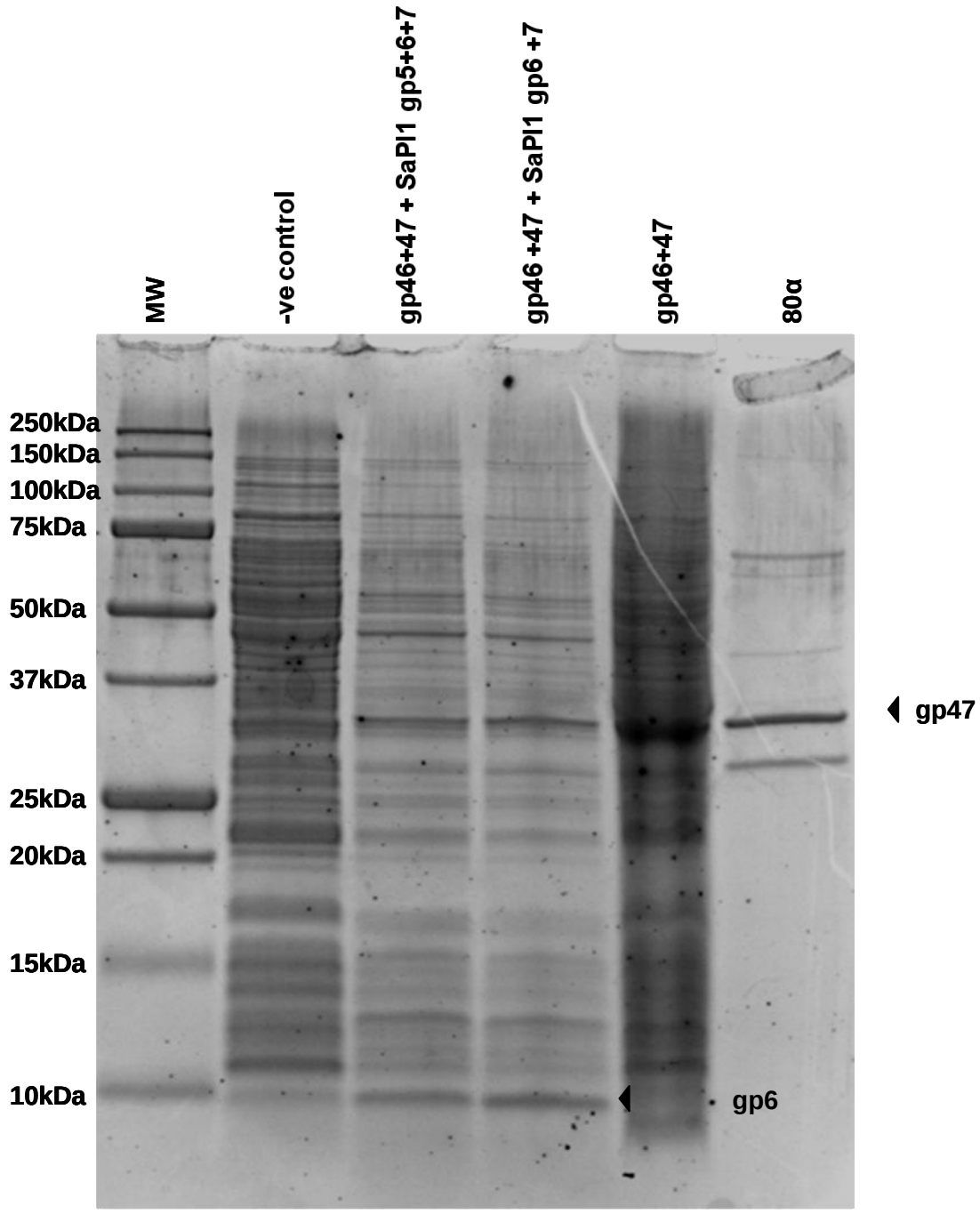
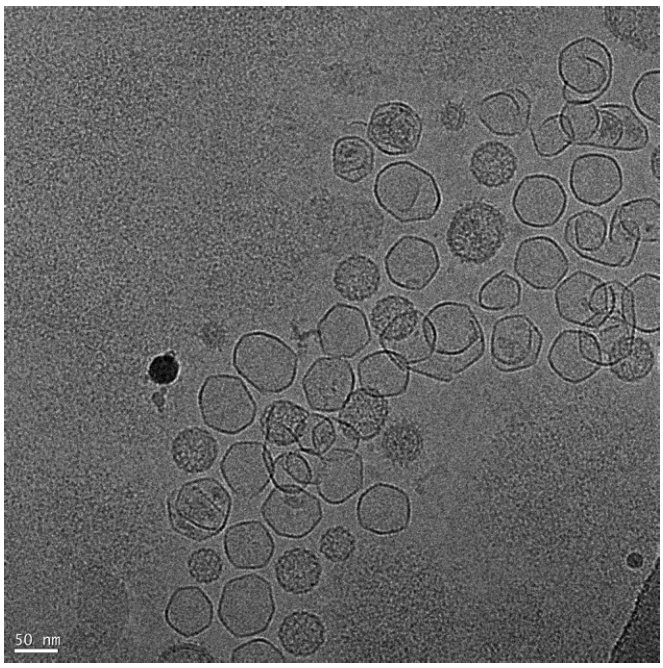

MW
-ve control
gp46+47 + SaPI1 gp5+6+7
gp46 +47 + SaPI1 gp6 +7
gp46+47
80α
250kDa
150kDa
100kDa
75kDa
50kDa
37kDa
25kDa
20kDa
15kDa
10kDa
gp47
gp6

Figure 37. Structures formed from coexpression of 80α major capsid and scaffold along with SaPI1 genes 6 and 7.

Particles were obtained by coexpressing cloned major capsid, scaffold SaPI1 genes 6 and 7(ST118). These particles were purified on CsCl and subsequently on sucrose gradients.
Cryo-EM was done by standard methods: 3 µL of sample was applied to non-glow-discharged C-flat holey film (Electron Microscopy Sciences, Hatfield, PA), blotted briefly before plunging into liquid ethane, and transferred to a Gatan 626 cryo-sample holder. Samples were observed in an FEI Tecnai F20 electron microscope operated at 200 kV, and images were captured on Kodak SO-163 film at a magnification of 62,000×.


50 nm

absence of P22 scaffolding protein results in formation of small (T=4) procapsids and a lot of spiral structures that cannot form closed procapsid like structures. Similar aberrant structures or small (T=4) procapsids were not observed when we expressed 80α gp47 alone. This failure to observe any kind of assembly may be attributed to the stability of the procapsid like structures. Though 80α and P22 major capsid proteins share a HK97 like fold, P22 has an additional telokine-like domain that has been implicated in stabilizing capsids (Parent *et al.*, 2010). However, we also cannot rule out the possibility that low expression of gp47 from plasmid pPD20 led to the observed lack of procapsids, although we would have expected to observe at least a few, discrete procapsid like structures even in this case. It is also possible that the apparent low expression of gp47 could be attributed to its high turnover in the absence of stable procapsid like structures. Further studies are needed to dissect which is the cause and which is the effect. However, it still can be safely concluded that the minimum subset of proteins required for formation of T=7 procapsid-like structures with a high fidelity consists of scaffold and major capsid proteins and that addition of SaPI1 gp6 and gp7 to this subset leads to formation of smaller (T=4) procapsids.

Chapter 7

Discussion

This work was designed to study the structural aspects of staphylococcal pathogenicity island SaPI1 transducing particles and its inter-relationship with helper phage 80α. SaPI1 exploits helper phage 80α functions in a highly efficient and specific manner for its own mobilization. The horizontal dissemination of SaPIs is not limited to intraspecies transfer, but has recently been shown capable of being intergeneric (Maiques *et al.*, 2007). Since staphylococcal PRCIs encode several known virulence factors, this transfer has wider implications in both health care and the food industry.

One striking aspect of SaPI mobilization is the capsid size redirection during the helper phage capsid assembly process, producing small sized virions that are large enough to package the entire SaPI genome, but not that of the helper phage. At the outset, we considered three possible ways by which morphologically distinct SaPI1 particles could be formed: (1) SaPI1 encodes its own alternative coat proteins; (2) phage 80α encoded capsid protein is alternatively processed to form the smaller capsids; or (3) capsid assembly is redirected by use of a different SaPI-encoded scaffold. Previous studies showed that SaPI virions consisted entirely of helper phage encoded proteins and that these proteins were identical (Tallent *et al.*, 2007), ruling out the first two possibilities. This argued for involvement of an alternate scaffold in capsid size redirection. Since scaffolding proteins are transiently involved in the self-assembly

of a procapsid, this hypothesis was consistent with the observations from the earlier studies.

The redirection of capsid size exhibited by SaPI1 is reminiscent of the exploitation of helper phage P2 by satellite phage P4 in *E. coli,* during which P4 directs formation of smaller capsids from P2 encoded structural proteins and thus favors its own encapsidation. The P4 capsid (T=4) is roughly one-third the volume of P2 capsid (T=7), commensurate with its genome size. Size determination of P4 capsids is carried out by a P4-encoded protein, Sid. This protein is transiently involved as an external scaffold during P4 procapsid assembly and allows only the formation of smaller capsids and not the larger ones. The P2-P4 paradigm of capsid size redirection provided an initial model on which we based our investigation of capsid size redirection in the 80α-SaPI1 system.

These studies have addressed questions related to the structures of helper phage capsids and SaPI1 virions, have identified SaPI1 proteins responsible for redirecting the capsid assembly process to assemble smaller sized SaPI1 virions, and have provided initial framework for understanding the self-assembly process. Below is a summary of our findings.

Role of scaffolding proteins in self-assembly of a procapsid

Most double stranded phages assemble their icosahedral shells from multiple copies of one capsid protein. The first stage in the assembly process is formation of an empty shell, known as a procapsid. However, formation of the hollow procapsid shell poses a significant challenge that requires anisotropic, directional interactions between subunits in space and time. The capsid protein by itself does not ensure correct

geometry of the capsid shell, but requires a scaffolding function for timely and proper procapsid assembly (Dokland, 1999; Fane *et al.*, 2003). These critical functions, in many phages, are performed by a separate dedicated scaffolding protein. Less frequently, as is seen in HK97 and T5, scaffold domains are embedded within the capsid proteins (Duda *et al.*, 1995a; Effantin *et al.*, 2006; Huet *et al.*, 2010). The rate of self-assembly is much slower in the absence of scaffolding proteins. In phage P22, for example, it has been shown that the lack of scaffolding protein decreases the rate of assembly by about tenfold (Casjens *et al.*, 1974).

Not only does self-assembly proceed at a slower rate in the absence of scaffold, but the number of aberrant structures increases. For example, mutation in the scaffolding protein of Φ29 results in assembly of isometric particles instead of prolate ones (Aksyuk *et al.*, 2011). Capsid proteins of T4 and phage λ assemble into cylindrical structures rather than icosahedral procapsids in the absence of their respective scaffolding proteins (Kellenberger,1980; Kellenberger,1990).

Scaffolding proteins have been shown to serve three important functions in the process of self-assembly. First, they facilitate the nucleation of assembly and thereby increase the rate of the self-assembly process (Caspar,1980; Erickson *et al.*, 1981). Second, they help incorporate other essential proteins such as the portal proteins into the procapsid (Earnshaw *et al.*, 1978; Moore *et al.*, 2002; Huffman *et al.*, 2008; Yang *et al.*, 2009a; Yang *et al.*, 2009b). Finally, they ensure proper geometry of procapsids by suppressing formation of aberrant structures during the self-assembly process (Salunke *et al.*, 1989; Thuman-Commike *et al.*, 1998).

Scaffolding proteins have been arbitrarily classified into three groups, based on their location and levels of complexity. Group 1 includes the icosahedrally ordered external scaffolding proteins of phage P4 and ΦX174-related phages. Group 2 includes the internal scaffolding proteins like those found in phages P22, and λ and in herpes virus. Group 3 includes proteins from phages with the multiple internal scaffolding proteins, such as those in the T4-like phages (Dokland,1999).

Structural studies of external scaffolding proteins have been performed for two bacteriophages, namely the ssDNA phage ΦX174 and *E. coli* dsDNA satellite phage P4. The external scaffolds of both these phages are predominantly α-helical, extended proteins that interact with the head proteins and form an external lattice around their respective procapsids. The external scaffold protein Sid of phage P4 is discussed here in some detail, since it provides the only characterized example of capsid size redirection. Phage P2 encodes a major capsid protein (gpN) and an internal scaffolding protein (gpO) that normally assemble into a large (T=7) P2 procapsid. The satellite phage P4 encoded external scaffolding protein, Sid, overrides the effect of gpO and redirects the self-assembly process to form smaller (T=4) procapsids.

Genetic studies suggest that Sid interacts only with gpN, since P2 mutants resistant to the action of Sid, known as *sir* mutants (Sid responsiveness), have been isolated only in gene *N* (Six *et al.*, 1991). These mutations clustered to the middle of gene *N,* suggesting that either this region of gpN was responsible for gpN-Sid interactions or it acted as a “hinge” that provided flexibility to gpN to assume the conformation required for the T=4 P4 capsid. In an attempt to dissect the above two possibilities, extragenic second site suppressors of *sir* mutations were isolated (Kim *et*

al., 2001). The mutations in these *nms* mutants (N mutation sensitive) map to the C-terminus of Sid and are thought to create stronger Sid-Sid interactions that can overcome the effects of the original *sir* mutations. Structural studies with the Sid protein showed that it forms a dodecahedral lattice over the T=4 procapsid (Marvik *et al.*, 1994b; Marvik *et al.*, 1995; Wang *et al.*, 2000). The outline of the lattice makes twelve large pentagons around the five-fold axis of symmetry that is occupied by the gpN pentamers. These lattices also bifurcate the 30 gpN hexamers and make trimeric Sid-Sid-Sid interactions between two gpN hexamers (Wang *et al.*, 2000; Dokland *et al.*, 2002). Thus the role of gpSid is to physically limit the formation of a large capsid by building a cage around it. It is noteworthy that gpO is required for assembly of viable P4 particles in vivo (Christie *et al.*, 1990) since it is required for the processing of gpN.

SaPI1 gp6 and gp7 are the capsid size redirecting factors

In order to identify SaPI proteins that could function as alternate scaffold proteins, we analyzed and compared the protein composition of phage 80α and SaPI1 capsids and procapsids by SDS-PAGE and mass spectrometry. The most significant findings from these experiments are (1) the presence of phage 80α encoded scaffold gp46 in both 80α and SaPI1 procapsids, but not in their corresponding mature virions, and (2) the presence of two SaPI1 proteins, gp6 and gp7, in SaPI1 procapsids but not in SaPI1 virions. gp6 and gp7 homologs of SaPIbov1 had earlier been implicated in formation of small capsids (Ubeda *et al.*, 2007). SaPI1 gp6 and gp7 are both abundant in SaPI1 procapsids concentrated on CsCl gradients, although gp7 appears to be greatly diminished in sucrose gradient purified particles, suggesting a loose association.

The presence of both these proteins in SaPI1 procapsids strongly suggested that they are involved in formation of small capsids.

SaPI1 gp6 and gp7 are sufficient and necessary for formation of small capsids

We performed loss-of-function mutation studies in SaPI1 by deleting orf6 and orf7 and observed their phenotypes by electron microscopy and by agarose gel electrophoresis of encapsidated DNA to confirm that these proteins are necessary for capsid size redirection. Furthermore, in order to confirm that these proteins are sufficient for capsid size redirection, we also constructed transcriptional fusions by inserting SaPI1 orfs 6 and 7 alone and together into an 80α prophage genome. We performed growth assays, cryo-EM analysis and Southern blots of the encapsidated DNA to assess whether these proteins are necessary and sufficient for capsid size redirection. We discovered that both of these proteins are required to redirect capsid size, and that no additional SaPI proteins are necessary.

SaPI1 size redirecting factors supplement 80α scaffold in formation of small capsids

In order to investigate the role of 80α encoded scaffold in formation of small capsids, we knocked out gp46 and observed its phenotype by performing biological assays and electron microscopy. Induction of *S. aureus* strain ST51 [RN4220 SaPI1 *tst::tetM*(80α Δ*orf46*)] and ST91 [RN4220(80α Δ*orf46*)] failed to produce viable phages or SaPI1 transducing particles. Also, electron microscopic examination of particles obtained from ST51 revealed just a few aberrant shells and some small procapsids, while lysate from ST91 did not show the presence of any kind of procapsid-like particles. Thus 80α encoded scaffold is required not only for formation of large capsids

but also the small sized SaPI capsids. The abundance of SaPI encoded proteins gp6 and gp7 in SaPI1 procapsids but not in SaPI1 virions indicated that they might act transiently to redirect capsid formation and that they act in conjunction with 80α encoded scaffold to form small capsids rather than replacing it.

80α major capsid and scaffold are N-terminally processed

Our initial experiments with in vivo expression of head proteins in *E. coli* revealed that the masses of both 80α scaffold and major capsid proteins were different from those of virions. SDS-PAGE analysis showed that both scaffold and capsid proteins from native procapsids migrated faster than ones overexpressed in *E. coli*. The faster migration rate of scaffold and major capsid proteins from native procapsids indicated that these proteins were post-translationally modified. Analysis by electrospray ionization (ESI) mass spectroscopy showed that both major capsid and scaffold proteins in 80α and SaPI1 procapsids were found to be cleaved at their N-terminus, within the consensus sequence KLKxNLQxF*A, where * denotes the cleavage site. The significant similarity at the N-terminus of these proteins strongly suggested that the proteins are cleaved by the same protease. There was also no difference in the masses of the two proteins between 80α and SaPI1, ruling out the possibility of differential processing of major capsid in SaPI1 capsids.

Major capsid and scaffold proteins are cleaved by a host protease

Cleavage of structural proteins is a common feature among double stranded DNA bacteriophages. Bacteriophages P2, T4 and HK97 are examples of double stranded bacteriophages that exhibit cleavage of the major capsid protein. Cleavage in

these phages occurs after procapsid assembly and is thought to be an important control point for the expansion of procapsids(Marvik *et al.*, 1994a; Conway *et al.*, 1995). The cleavage in these phages is carried out by a phage protease that is usually located immediately upstream of scaffold gene, in the same operon. Sometimes the protease function is embedded in the scaffold itself as is the case for phage T4 (van Driel *et al.*, 1980). Based on these observations from other phages, we hypothesized that 80α orf44, which lies immediately upstream of the 80α scaffold gene, could be a putative protease. gp44 was also found in 80α and SaPI1 procapsids and mature 80α virions indicating that orf44 was a real gene and hence was the most likely candidate for a phage encoded protease. In order to test our hypothesis, we conducted SDS-PAGE analysis, cryo-electron microscopy and biological assays with an orf44 deletion mutant. Our observations strongly suggested that gp44 is not a protease. A search in the reference protein database (NCBI) using PSI-BLAST (Position-specific iterated BLAST) (Altschul *et al.*, 1997) with the gp44 amino acid sequence as a query showed that this protein had homologs in other Staphylococcal Siphoviridae and a slight similarity to phage SPP1 gp6 protein. SPP1 gp6 is required for efficient infection of B.subtilis and has been demonstrated to play a role in DNA ejection (Vinga *et al.*, 2006). Deletion of the gene encoding SPP1 gp6 causes premature DNA ejection and significant loss of infectivity. Thus based on all these data, we suggest that gp44 is a “neck” protein involved in head-tail connection and/or DNA stability.

When major capsid protein was expressed in *S. aureus* and compared to major capsid protein from native 80α particles by SDS-PAGE, their migration rates were identical. Subsequently, masses of both these proteins were found to be identical by

mass spectrometry (Dr Terje Dokland personal communication). These observations strongly suggest that a host protease is involved in N-terminal processing of 80α head proteins. This is unusual for prohead processing because of its critical role in capsid maturation. All phages studied thus far that have prohead processing encode their own proteases. Though using a host protease might be a good tactic to reduce the coding capacity of its genome, relying on a host for such a critical process could prove disastrous for the phage unless the protease is also essential to the host.

If indeed a host protease is responsible for cleaving major capsid and scaffold proteins then it should also have a similar function in *S. aureus*. In order to find targets of this putative protease, we performed a BLAST search against the *S. aureus* genome using the N-terminus of major capsid and scaffold protein sequences as a query. This search identified a single protein with sequence similarity at the N-terminus, staphylococcal ribosomal protein L27 (NCBI Reference Sequence: YP_500260.1).

L27 is a component of the ribosomal large (50S) subunit. Genetic experiments in *E. coli* have shown its role in peptidyl transferase activity. L27 proteins from the Gram-positive bacteria *S. aureus* and *B. subtilis* are homologous to the *E. coli* L27 except that they have an extra sequence at their N-terminus that is similar to the N-terminal sequence of the 80α scaffold and major capsid proteins. Furthermore, the active from of *E. coli* L27 lacks the formyl methionine and starts with an alanine as would the staphylococcal L27 if its N-terminal region was cleaved. These observations strongly suggest that staphylococcal L27 is N terminally processed between Phe and Ala to yield the active form.

If the staphylococcal ribosomal protein L27 is cleaved at the N-terminus, then this processing must have an important role in the biological process. Microbes have evolved to maximize efficiency, and encoding useless regions would go against this paradigm. It is common for proteins to have a signal sequence in the N-terminal region and this is generally true for proteins that are secreted or targeted to the cellular membranes. However, L27 is a cytoplasmic protein and hence neither of the roles for its N-terminal region seem significant in its case. Of the several possible roles for the N-terminus of S. aureus L27, two seem attractive. First, is the N-terminus of *S. aureus* an intramolecular chaperone? Some proteins, such as subtsilin, use their N-terminus as an intramolecular chaperone. These proteins have an N-terminal region about 70 to 200 amino acids long that assists the propetide in folding properly. However, this possibility also seems to be unlikely due to the fact that subtsilin-like proteins have an autoproteolytic activity while the staphylococcal L27, based on our observations with major capsid and scaffold proteins, would not have such an activity. Furthermore, the N-terminal region is too short and has a propensity to form only a single α-helix that might not be sufficient to assist folding of the remaining primary structure. Second, does the N-terminus have a role in the ribosomal protein autoregulatory network? Ribosomal protein synthesis is translationally regulated so that the ribosomal proteins are synthesized in stoichiometric proportions. This control is important if the bacteria has to cope with varying environmental conditions. Translational control by feedback-regulation of ribosomal protein synthesis has been extensively studied in *E. coli*. However, these mechanisms remain poorly understood in Gram positive bacteria. It is possible that the truncated N-terminal peptide of the staphylococcal L27 has some role

in this regulatory network. High levels of these cleaved peptides or accumulation of the uncleaved form could provide some level of feedback and signal the cells to stop ribosomal protein synthesis. Perhaps, the phage could be possibly exploiting this signaling mechanism so that while it is multiplying in the cell, resources are not diverted towards the host functions. Clearly more studies would be required in identifying the role of these cleaved fragments.

The morphology of helper phage particles is distinct from SaPI1 particles

Our cryo-electron microscopic studies of helper phage 80α and SaPI1 particles show that both these particles have a 190 nm long flexuous tail capped with a baseplate containing six tail fibers. The phage capsids are 63 nm in diameter and are isometric and icosahedral in shape. The SaPI1 capsids are also isometric and icosahedral in shape but only 46 nm in diameter. Three dimensional (3-D) reconstruction with cryo-EM images of phage 80α and SaPI1 capsids and procapsids reveals that they have an icosahedral symmetry corresponding to triangulation numbers T=7*l* and T=4, respectively.

Furthermore, the HK97-like capsid fold that has been found in all *Caudovirales* studied thus far, and even in herpes virus (Bamford *et al.*, 2005; Johnson *et al.*, 2007) is also seen in the 80α gp47 capsid protein. Like other capsid proteins of T=7 phages, 80α major capsid protein exists in 7 conformational states with minor differences in their tertiary structure. Though 80α capsid protein has a HK97-like fold, there are major differences between gp47 and HK97 gp5 throughout the structure. The biggest difference is in the elongated P loop of gp47, which was found to form a trifoliate β hairpin at the threefold inter-capsomer contacts in the 80α icosahedral shell (Spilman *et*

al., 2011). This trifoliate density is reflected in a clustering of strong density in the reconstruction, suggesting that this region is critical for capsid stability. This region has been found to be critical in providing capsid stability in phage λ and P22 as well. For example, in phage λ there is an additional decoration protein (gpD) at this location that strengthens the three fold intercapsomer interaction (Lander *et al.*, 2008). P22 has an extended P-loop similar to that of gp47 and an additional telokine domain that is involved in capsid stability (Parent *et al.*, 2010). In HK97, residues involved in cross linking are located in this region (Wikoff *et al.*, 2000). The other important observation is the presence of an α helix in the procapsid that is not found in mature capsid. This suggests that this arm is involved in the maturation process and possibly interacts with the scaffold protein (Spilman *et al.*, 2011).

The mechanism of conformational change during maturation in 80a major capsid protein differs from that of HK97. Capsid maturation in HK97 gp5 is caused by a rotation of the A domain by 39° relative to the P domain. The most striking conformational change in the 80α capsid protein during capsid maturation occurs in the spine helix (α3), due to proline at 132 that produces a kink in the helix. The ψ angle of Pro132 switches from + 135° to – 45° during the maturation of the procapsid. This change in the ψ angle causes a rotation of the N-terminal half of α3 relative to the rest of the helix and this change is propagated through α2 to the whole P domain, including the P loop. However, the position of the A domain relative to the P domain remains essentially the same. Thus in spite of having similar tertiary folds, these capsid proteins employ different mechanisms to achieve the final stable conformation pointing to the versatility of the maturation process. Further studies need to be directed at understanding how

80α gp47 can assume conformations of a T=4 state. These studies would help in understanding the plasticity of 80α gp47 and how it can form stable structures that have an altogether different triangulation number.

SaPI1 procapsid assembly is mediated by an internal scaffold

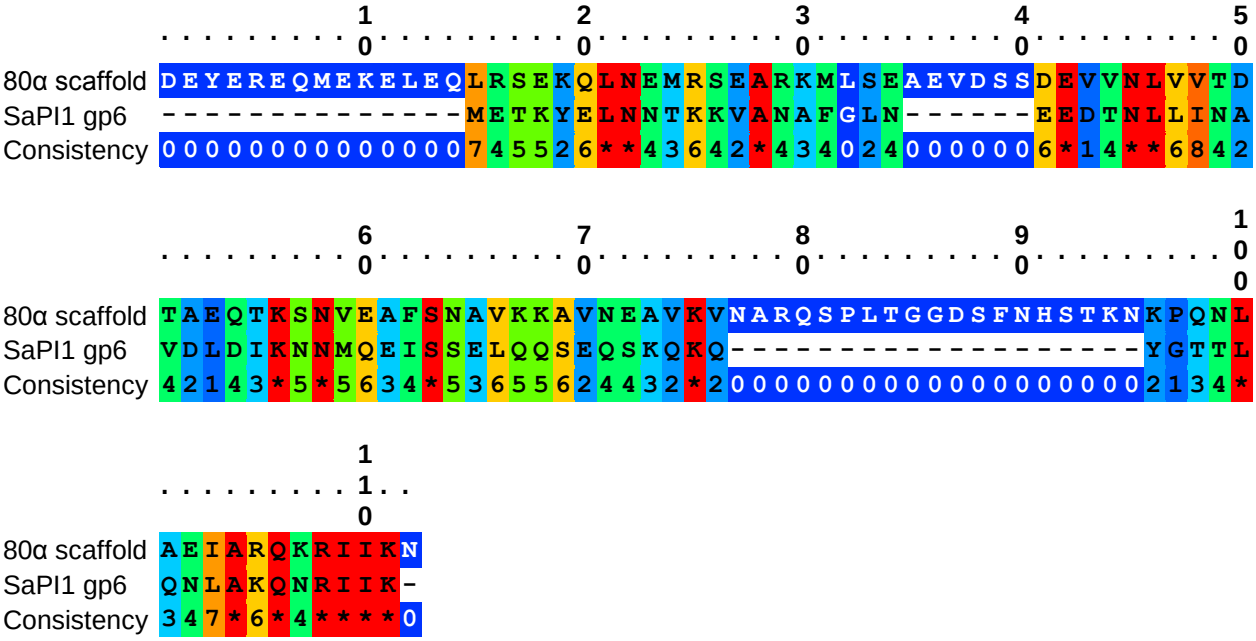
The phage 80α and SaPI1 procapsids, similar to the capsid precursors of other *Caudovirales*, are smaller and more rounded than their mature capsids, These procapsids have icosahedral symmetries corresponding to their mature capsids. However, their walls are thicker than of mature capsid with hexamers and pentamers that protrude more distinctly than in mature capsids. The capsid protein subunits are more closely packed than in mature capsids.

Most importantly, the external surface of SaPI1 procapsids was as smooth as that of 80α procapsids and lacked any additional features that could be attributed to a Sid-like external scaffold on the exterior shell of the SaPI1 procapsid, similar to those seen on P4 procapsids. The interior of SaPI1 procapsids, however, unlike the 80α procapsids, has 120 nm long finger-like, well defined densities that protrude from the inner surface to the center of the procapsid. These observations marked a major deviation from our adopted P2-P4 model because it suggested that the SaPI1 procapsid assembly is not mediated by an external scaffold, as in P4 procapsids, but is redirected by alternative internal scaffolds. Indeed both gp6 and gp7 have propensity to form α-helical structures, which are a hallmark of internal scaffold proteins. Interestingly, SaPI1 gp6 and 80α scaffold protein share a sequence similarity in the C-terminal region (Fig 38). The C-terminus has been found to be important in other phages as well. Recent

Figure 38. Sequence alignment of C-terminal sequences of 80α scaffold and SaPI1 gp6

Shown is a sequence alignment of the C-terminal sequence of 80α scaffold and SaPI1 gp6
The alignment was performed using an online multiple sequence alignment program: Praline, IBIVU Vrije Universiteit Amsterdam

The conservation scoring is performed by PRALINE. The scoring scheme works from 0 for the least conserved alignment position, up to 10 for the most conserved alignment position.
The color assignments are:


80α scaffold
SaPI1 gp6
Consistency
DEYEREQMEKELEQLRSEKQLNEMRSEARKMLSEAEVDSSDEVVNLVVTD
METKYELNNTKKVANAFGLNEEDTNLLINA
0000000000000007455265**43642*43402400000006*14**6842
TAEQTKSNVEAFSNAVKKAVNEAVKVNARQSPLTGGDSFNHSTKNKPQNL
VDLDIKNNMQEISSELQQSEQSKQKQYGTTL
4214356*5*56345*536556244324*200000000000000000000213 4*
AEIARQKRIIKN
QNLAKQNRIIK-
3476*6*4****0

studies with the P22 scaffold have shown that the interaction of the P22 encoded scaffold with its coat protein occurs through the C-terminus (Chen *et al.*, 2011). Most 80α rescue mutants obtained from the plaques of 80α insertion gp6 and gp7 mutant map to the C-terminal region of SaPI1 orf6 (personal communication Dr. Gail E. Christie). These mutations suggest that the C-terminal region of gp6 is very critical to capsid size redirection and since there is remarkable similarity between the C-termini of 80α scaffold and SaPI1 gp6, both these proteins are likely interacting with the capsid subunits through this region.

The role of SaPI1 gp7 in capsid size redirection is not clear at present. This protein along with gp6 is clearly required for formation of smaller sized capsids as suggested by our experiments. Interestingly, in our experiments with insertion and deletion of SaPI1 size redirection factors, when gp7 was present alone with 80α morphogenetic proteins, a few smaller capsids were also produced. Experiments with P22, in which the levels of scaffold were modulated also produced a few smaller capsids along with the larger ones (Thuman-Commike *et al* 1998). Possibly gp7 interacts with 80α scaffold and thereby makes its unavailable for capsid assembly. This results in capsid assembly with lower fidelity and yields large and small capsids. When gp6 is also present in this system, it does not face any competition from 80α scaffold and can interact with the 80α capsid protein to predominantly produce small capsids. Further experiments to establish the role of SaPI1 gp7 are required.

<u>The intriguing phenomenon of interference</u>

The 80α–SaPI1 relationship causes interference with the propagation of 80α in presence of SaPI1. This interference is so drastic that the 80α burst size is reduced by about two orders of magnitude. Two possible explanations for this phenomenon of interference are: altered packaging specificity due to the SaPI1 encoded small terminase, and formation of small capsids that physically limit encapsidation of a full length 80α genome. Earlier studies in our lab demonstrated that deletion of SaPI1 *terS* decreased SaPI1 transduction by at least three orders of magnitude, but did not relieve interference in such mutants. However, formation of small capsids would still be normal and might mask an effect of packaging redirection. We expected that small capsid formation would interfere with 80α propagation, due to the inability to package a complete 80α genome. This was indeed the case when SaPI1 genes 6 and 7 were inserted into the 80α capsid gene cluster.

Surprisingly, however, deleting orf6 or 7 did not relieve this interference even though capsids were now large. These results suggest that the interference of helper phage 80α growth is a multifactorial event and other genes, such SaPI1 encoded *terS* in combination with capsid size determination genes, are involved. Further studies aimed at deleting these genes in various combinations and identification of putative involvement of other SaPI1 genes would be needed to explain this effect.

Based on this study, the current model of capsid size redirection (Fig 39) in the 80α-SaPI1 system is follows: capsid assembly is initiated with multiple copies of the major capsid protein to form a procapsid. 80α encoded scaffold is needed during the self-assembly to yield 80α sized procapsids. The portal complex, which is important for

Figure 39. Current model of capsid size redirection

terS
terL
Portal
Minor Capsid
Scaffold
Major Capsid
Major capsid
Scaffold
Minor capsid
Portal
gp6
gp7?

providing a route for the entry and exit is also incorporated during this stage as is the minor capsid protein, which is possibly a neck protein present in low amounts. This transient intermediate undergoes morphological transformation with the removal of scaffold and DNA packaging to yield a mature 80α capsid. However, in the presence of SaPI1 encoded gp6 and gp7 capsid assembly is redirected to form smaller sized procapsids. gp6 supplements the role of 80α scaffold as an internal scaffold, possibly by interacting through its C-terminus with the major capsid protein. The role of gp7 is not clear at present and we postulate that it competitively inhibits association of 80α scaffold with major capsid proteins. These small procapsids mature with the release of gp6, gp7 and 80α scaffold to form SaPI1 capsids.

References

From the centers for disease control. *Staphylococcus aureus* resistant to vancomycin--United States, 2002. (2002) *JAMA* **288:** 824-825.

New and nonofficial drugs: Vancomycin hydrochloride. (1959) *J Am Med Assoc* **170:** 810-811.

Abinun M., and Cant A.J. (1994) Toxic shock syndrome toxin-secreting *Staphylococcus aureus* in kawasaki syndrome. *Lancet* **343:** 300.

Ackermann H. (2001) Frequency of morphological phage descriptions in the year 2000. *Arch Virol* **146:** 843-857.

Akada R. (1994) Quick-check method to test the size of *Escherichia coli* plasmids. *BioTechniques* **17:** 58.

Aksyuk A.A., and Rossmann M.G. (2011) Bacteriophage assembly. *Viruses* **3:** 172-203.

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., and Lipman D.J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389-3402.

Aman M.J., Karauzum H., Bowden M.G., and Nguyen T.L. (2010) Structural model of the pre-pore ring-like structure of panton-valentine leukocidin: Providing dimensionality to biophysical and mutational data. *J Biomol Struct Dyn* **28:** 1-12.

Arnaud M., Chastanet A., and Debarbouille M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70:** 6887-6891.

Arnold R.J., and Reilly J.P. (1999) Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry. *Anal Biochem* **269:** 105-112.

Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K., Oguchi A.*, et al.* (2002) Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359:** 1819-1827.

Babe L.M., and Craik C.S. (1997) Viral proteases: Evolution of diverse structural motifs to optimize function. *Cell* **91:** 427-430.

Baker T.S., Olson N.H., and Fuller S.D. (1999) Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol Mol Biol Rev* **63:** 862-922.

Bamford D.H., Grimes J.M., and Stuart D.I. (2005) What does structure tell us about virus evolution? *Curr Opin Struct Biol* **15:** 655-663.

Barrett K., Calendar R., Gibbs W., Goldstein R.N., Lindqvist B., and Six E. (1973) Helper-dependent bacteriophage P4: A model satellite virus and its implications for animal virology. *Prog Med Virol* **15:309-30:** 309-330.

Bateman A., and Rawlings N.D. (2003) The CHAP domain: A large family of amidases including GSP amidase and peptidoglycan hydrolases. *Trends Biochem Sci* **28:** 234-237.

Becker S.C., Foster-Frey J., Stodola A.J., Anacker D., and Donovan D.M. (2009) Differentially conserved staphylococcal SH3b_5 cell wall binding domains confer increased staphylolytic and streptolytic activity to a streptococcal prophage endolysin domain. *Gene* **443:** 32-41.

Bekeredjian-Ding I., Inamura S., Giese T., Moll H., Endres S., Sing A.*, et al.* (2007) *Staphylococcus aureus* protein A triggers T cell-independent B cell proliferation by sensitizing B cells for TLR2 ligands. *J Immunol* **178:** 2803-2812.

Blomster-Hautamaa D.A., Kreiswirth B.N., Kornblum J.S., Novick R.P., and Schlievert P.M. (1986) The nucleotide and partial amino acid sequence of toxic shock syndrome toxin-1. *J Biol Chem* **261:** 15783-15786.

Bohach G.A. (2006) *Staphylococcus aureus* exotoxins. In Gram-positive pathogens.2nd Ed (2006). **ISBN-10:** 1555813437 American society for Microbiology, pp. 464-477.

Bohach G.A., Fast D.J., Nelson R.D., and Schlievert P.M. (1990) Staphylococcal and streptococcal pyrogenic toxins involved in toxic shock syndrome and related illnesses. *Crit Rev Microbiol* **17:** 251-272.

Botstein D. (1980) A theory of modular evolution for bacteriophages. *Ann N Y Acad Sci* **354:** 484-490.

Camacho A., Jimenez F., Vinuela E., and Salas M. (1979) Order of assembly of the lower collar and the tail proteins of *Bacillus subtilis* bacteriophage phi 29. *J Virol* **29:** 540-545.

Casjens S., and King J. (1974) P22 morphogenesis. I: Catalytic scaffolding protein in capsid assembly. *J Supramol Struct* **2:** 202-224.

Caspar D.L. (1980) Movement and self-control in protein assemblies. quasi-equivalence revisited. *Biophys J* **32:** 103-138.

Caspar D.L., and Klug A. (1962) Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* **27:** 1-24.

Centers for Disease Control and Prevention (CDC). (2002) *Staphylococcus aureus* resistant to vancomycin--United States, 2002. *MMWR Morb Mortal Wkly Rep* **51:** 565-567.

Chang J.R., Spilman M.S., Rodenburg C.M., and Dokland T. (2009) Functional domains of the bacteriophage P2 scaffolding protein: Identification of residues involved in assembly and protease activity. *Virology* **384:** 144-150.

Chang J.R., Poliakov A., Prevelige P.E., Mobley J.A., and Dokland T. (2008) Incorporation of scaffolding protein gpO in bacteriophages P2 and P4. *Virology* **370:** 352-361.

Chang S.Y., McGary E.C., and Chang S. (1989) Methionine aminopeptidase gene of *Escherichia coli* is essential for cell growth. *J Bacteriol* **171:** 4071-4072.

Chen D.H., Baker M.L., Hryc C.F., Dimaio F., Jakana J., Wu W.*, et al.* (2011) Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proc Natl Acad Sci U S A* 108(4): 1355–1360.

Cheung A.L., Nishina K.A., Trotonda M.P., and Tamber S. (2008) The SarA protein family of *Staphylococcus aureus*. *Int J Biochem Cell Biol* **40:** 355-361.

Christie G.E., and Calendar R. (1990) Interactions between satellite bacteriophage P4 and its helpers. *Annu Rev Genet* **24:** 465-490.

Christie G.E., Matthews A.M., King D.G., Lane K.D., Olivarez N.P., Tallent S.M.*, et al.* (2010) The complete genomes of *Staphylococcus aureus* bacteriophages 80 and 80alpha--implications for the specificity of SaPI mobilization. *Virology* **407:** 381-390.

Clarke S.R., and Foster S.J. (2006) Surface adhesins of *Staphylococcus aureus*. *Adv Microb Physiol* **51:** 187-224.

Clauditz A., Resch A., Wieland K.P., Peschel A., and Gotz F. (2006) Staphyloxanthin plays a role in the fitness of *Staphylococcus aureus* and its ability to cope with oxidative stress. *Infect Immun* **74:** 4950-4953.

Colin D.A., Mazurier I., Sire S., and Finck-Barbancon V. (1994) Interaction of the two components of leukocidin from *Staphylococcus aureus* with human polymorphonuclear leukocyte membranes: Sequential binding and subsequent activation. *Infect Immun* **62:** 3184-3188.

Conway J.F., Duda R.L., Cheng N., Hendrix R.W., and Steven A.C. (1995) Proteolytic and conformational control of virus capsid maturation: The bacteriophage HK97 system. *J Mol Biol* **253:** 86-99.

Cooney J., Kienle Z., Foster T.J., and O'Toole P.W. (1993) The gamma-hemolysin locus of *Staphylococcus aureus* comprises three linked genes, two of which are identical to the genes for the F and S components of leukocidin. *Infect Immun* **61:** 768-771.

Cosgrove S.E., Sakoulas G., Perencevich E.N., Schwaber M.J., Karchmer A.W., and Carmeli Y. (2003) Comparison of mortality associated with methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* bacteremia: A meta-analysis. *Clin Infect Dis* **36:** 53-59.

Curtis N., Chan B., and Levin M. (1994) Toxic shock syndrome toxin-secreting *Staphylococcus aureus* in Kawasaki Syndrome. *Lancet* **343:** 299.

Deleo F.R., Otto M., Kreiswirth B.N., and Chambers H.F. (2010) Community-associated methicillin-resistant *Staphylococcus aureus*. *Lancet* **375:** 1557-1568.

D'Elia M.A., Pereira M.P., Chung Y.S., Zhao W., Chau A., Kenney T.J.*, et al.* (2006) Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. *J Bacteriol* **188:** 4183-4189.

Diana C., Deho G., Geisselsoder J., Tinelli L., and Goldstein R. (1978) Viral interference at the level of capsid size determination by satellite phage P4. *J Mol Biol* **126:** 433-445.

Diaz E., Lopez R., and Garcia J.L. (1990) Chimeric phage-bacterial enzymes: A clue to the modular evolution of genes. *Proc Natl Acad Sci U S A* **87:** 8125-8129.

Dinges M.M., Orwin P.M., and Schlievert P.M. (2000) Exotoxins of *Staphylococcus aureus*. *Clin Microbiol Rev* **13:** 16-34.

Dodd I.B., Shearwin K.E., and Egan J.B. (2005) Revisited gene regulation in bacteriophage lambda. *Curr Opin Genet Dev* **15:** 145-152.

Dokland T. (1999) Scaffolding proteins and their role in viral assembly. *Cell Mol Life Sci* **56:** 580-603.

Dokland T., and Murialdo H. (1993) Structural transitions during maturation of bacteriophage lambda capsids. *J Mol Biol* **233:** 682-694.

Dokland T., Wang S., and Lindqvist B.H. (2002) The structure of P4 procapsids produced by coexpression of capsid and external scaffolding proteins. *Virology* **298:** 224-231.

Dokland T., Lindqvist B.H., and Fuller S.D. (1992) Image reconstruction from cryo-electron micrographs reveals the morphopoietic mechanism in the P2-P4 bacteriophage system. *EMBO J* **11:** 839-846.

Duda R.L., Martincic K., and Hendrix R.W. (1995a) Genetic basis of bacteriophage HK97 prohead assembly. *J Mol Biol* **247:** 636-647.

Duda R.L., Martincic K., Xie Z., and Hendrix R.W. (1995b) Bacteriophage HK97 head assembly. *FEMS Microbiol Rev* **17:** 41-46.

Dutton A.A., and Elmes P.C. (1959) Vancomycin: Report on treatment of patients with severe staphylococcal infections. *Br Med J* **1:** 1144-1149.

Earnshaw W., and King J. (1978) Structure of phage P22 coat protein aggregates formed in the absence of the scaffolding protein. *J Mol Biol* **126:** 721-747.

Earnshaw W.C., and Casjens S.R. (1980) DNA packaging by the double-stranded DNA bacteriophages. *Cell* **21:** 319-331.

Effantin G., Figueroa-Bossi N., Schoehn G., Bossi L., and Conway J.F. (2010) The tripartite capsid gene of salmonella phage gifsy-2 yields a capsid assembly pathway engaging features from HK97 and lambda. *Virology* **402:** 355-365.

Effantin G., Boulanger P., Neumann E., Letellier L., and Conway J.F. (2006) Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. *J Mol Biol* **361:** 993-1002.

Erickson H.P., and Pantaloni D. (1981) The role of subunit entropy in cooperative assembly. Nucleation of microtubules and other two-dimensional polymers. *Biophys J* **34:** 293-309.

Fairbrother R.W., and Williams B.L. (1956) Two new antibiotics; antibacterial activity of novobiocin and vancomycin. *Lancet* **271:** 1177-1178.

Fane B.A., and Prevelige P.E.,Jr. (2003) Mechanism of scaffolding-assisted viral assembly. *Adv Protein Chem* **64:** 259-299.

Ferrer M.D., Quiles-Puchalt N., Harwich M.D., Tormo-Mas M.A., Campoy S., Barbe J.*, et al.* (2011) RinA controls phage-mediated packaging and transfer of virulence genes in Gram-positive bacteria. *Nucleic Acids Res doi: 10.1093/nar/gkr158*.

Foster T. (1996) Staphylococcus. In Medical Microbiology. Baron, S. (ed). Galveston (TX): The University of Texas Medical Branch at Galveston,Chapter 12 .

Foster T.J. (2005) Immune evasion by staphylococci *Nat Rev Microbiol* **3:** 948-958.

Fraser J.D., and Proft T. (2008) The bacterial superantigen and superantigen-like proteins. *Immunol Rev* **225:** 226-243.

Frost L.S., Leplae R., Summers A.O., and Toussaint A. (2005) Mobile genetic elements: The agents of open source evolution. *Nat Rev Microbiol* **3:** 722-732.

Fujisawa H., and Morita M. (1997) Phage DNA packaging. *Genes Cells* **2:** 537-545.

Gan L., Speir J.A., Conway J.F., Lander G., Cheng N., Firek B.A.*, et al.* (2006) Capsid conformational sampling in HK97 maturation visualized by X-ray crystallography and cryo-EM. *Structure* **14:** 1655-1665.

Garcia P., Garcia J.L., Garcia E., Sanchez-Puelles J.M., and Lopez R. (1990) Modular organization of the lytic enzymes of streptococcus pneumoniae and its bacteriophages. *Gene* **86:** 81-88.

Gertsman I., Gan L., Guttman M., Lee K., Speir J.A., Duda R.L.*, et al.* (2009) An unexpected twist in viral capsid maturation. *Nature* **458:** 646-650.

Giese B., Glowinski F., Paprotka K., Dittmann S., Steiner T., Sinha B., and Fraunholz M.J. (2011) Expression of delta-toxin by *Staphylococcus aureus* mediates escape from phago-endosomes of human epithelial and endothelial cells in the presence of beta-toxin. *Cell Microbiol* **13:** 316-329.

Giglione C., Boularot A., and Meinnel T. (2004) Protein N-terminal methionine excision. *Cell Mol Life Sci* **61:** 1455-1474.

Giraudo A.T., Cheung A.L., and Nagel R. (1997) The *sae* locus of *Staphylococcus aureus* controls exoprotein synthesis at the transcriptional level. *Arch Microbiol* **168:** 53-58.

Goerke C., Fluckiger U., Steinhuber A., Bisanzio V., Ulrich M., Bischoff M.*, et al.* (2005) Role of *Staphylococcus aureus* global regulators Sae and sigmaB in virulence gene expression during device-related infection. *Infect Immun* **73:** 3415-3421.

Goldberg M.M. (1937) A class of multi-symmetric polyhedra. *Tohoku Mathematical Journal* 104-108.

Gordon R.J., and Lowy F.D. (2008) Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. *Clinical Infectious Diseases* **46:** S350-S359.

Grundling A., Manson M.D., and Young R. (2001) Holins kill without warning. *Proc Natl Acad Sci U S A* **98:** 9348-9352.

Hacker J., Blum-Oehler G., Muhldorfer I., and Tschape H. (1997) Pathogenicity islands of virulent bacteria: Structure, function and impact on microbial evolution. *Mol Microbiol* **23:** 1089-1097.

Haggar A., Hussain M., Lonnies H., Herrmann M., Norrby-Teglund A., and Flock J.I. (2003) Extracellular adherence protein from *Staphylococcus aureus* enhances internalization into eukaryotic cells. *Infect Immun* **71:** 2310-2317.

Harraghy N., Hussain M., Haggar A., Chavakis T., Sinha B., Herrmann M., and Flock J.I. (2003) The adhesive and immunomodulating properties of the multifunctional *Staphylococcus aureus* protein Eap. *Microbiology* **149:** 2701-2707.

Harwich M.,Jr. (2009) Transcriptional profiling of staphylococcal bacteriophage 80 alpha and regulatory interactions with pathogenicity island SaPI1. PhD dissertation Virginia Commonwealth University Richmond VA : AAT3358185

Hauck C.R., and Ohlsen K. (2006) Sticky connections: Extracellular matrix protein recognition and integrin-mediated cellular invasion by *Staphylococcus aureus*. *Curr Opin Microbiol* **9:** 5-11.

Hendrix R.W. (2003) Bacteriophage genomics. *Curr Opin Microbiol* **6:** 506-511.

Hendrix R.W., and Duda R.L. (1998) Bacteriophage HK97 head assembly: A protein ballet. *Adv Virus Res* **50:** 235-288.

Horsburgh M.J., Ingham E., and Foster S.J. (2001) In *Staphylococcus aureus*, *fur* is an interactive regulator with PerR, contributes to virulence, and is necessary for oxidative stress resistance through positive regulation of catalase and iron homeostasis. *J Bacteriol* **183:** 468-475.

Horton R.M. (1995) PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. *Mol Biotechnol* **3:** 93-99.

Horton R.M., Cai Z.L., Ho S.N., and Pease L.R. (1990) Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. *BioTechniques* **8:** 528-535.

Huang R.K., Khayat R., Lee K.K., Gertsman I., Duda R.L., Hendrix R.W., and Johnson J.E. (2011) The prohead-I structure of bacteriophage HK97: Implications for scaffold-mediated control of particle assembly and maturation. *J Mol Biol* **408:** 541-554.

Huet A., Conway J.F., Letellier L., and Boulanger P. (2010) In vitro assembly of the T=13 procapsid of bacteriophage T5 with its scaffolding domain. *J Virol* **84:** 9350-9358.

Huffman J.B., Newcomb W.W., Brown J.C., and Homa F.L. (2008) Amino acids 143 to 150 of the herpes simplex virus type 1 scaffold protein are required for the formation of portal-containing capsids. *J Virol* **82:** 6778-6781.

Huseby M., Shi K., Brown C.K., Digre J., Mengistu F., Seo K.S.*, et al.* (2007) Structure and biological activities of beta toxin from *Staphylococcus aureus*. *J Bacteriol* **189:** 8719-8726.

Huseby M.J., Kruse A.C., Digre J., Kohler P.L., Vocke J.A., Mann E.E.*, et al.* (2010) Beta toxin catalyzes formation of nucleoprotein matrix in staphylococcal biofilms. *Proc Natl Acad Sci U S A* **107:** 14407-14412.

Iandolo J.J., Worrell V., Groicher K.H., Qian Y., Tian R., Kenton S.*, et al.* (2002) Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of *Staphylococcus aureus* 8325. *Gene* **289:** 109-118.

Inman R.B., Schnos M., Simon L.D., Six E.W., and Walker D.H.,Jr. (1971) Some morphological properties of P4 bacteriophage and P4 DNA. *Virology* **44:** 67-72.

James P., Quadroni M., Carafoli E., and Gonnet G. (1993) Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun* **195:** 58-64.

Johnson J.E., and Chiu W. (2007) DNA packaging and delivery machines in tailed bacteriophages. *Curr Opin Struct Biol* **17:** 237-243.

Jones D.T. (1999) GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. *J Mol Biol* **287:** 797-815.

Kanamaru S., Kondabagil K., Rossmann M.G., and Rao V.B. (2004) The functional domains of bacteriophage T4 terminase. *J Biol Chem* **279:** 40795-40801.

Kaneko J., Kimura T., Narita S., Tomita T., and Kamio Y. (1998) Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage phiPVL carrying panton-valentine leukocidin genes. *Gene* **215:** 57-67.

Kellenberger E. (1990) Form determination of the heads of bacteriophages. *Eur J Biochem* **190:** 233-248.

Kellenberger E. (1980) Control mechanisms in the morphogeneses of bacteriophage heads. *BioSystems* **12:** 201-223.

Kernodle D.S. (2007) Re: Is panton-valentine leukocidin the major virulence determinant in community-associated methicillin-resistant *Staphylococcus aureus* disease? *J Infect Dis* **195:** 1726.

Kikuchi Y., and King J. (1975a) Genetic control of bacteriophage T4 baseplate morphogenesis. III. formation of the central plug and overall assembly pathway. *J Mol Biol* **99:** 695-716.

Kikuchi Y., and King J. (1975b) Genetic control of bacteriophage T4 baseplate morphogenesis. II. mutants unable to form the central part of the baseplate. *J Mol Biol* **99:** 673-694.

Kikuchi Y., and King J. (1975c) Genetic control of bacteriophage T4 baseplate morphogenesis. I. sequential assembly of the major precursor, in vivo and in vitro. *J Mol Biol* **99:** 645-672.

Kim A.I., Adal K.A., and Schmitt S.K. (2003) *Staphylococcus aureus* bacteremia: Using echocardiography to guide length of therapy. *Cleve Clin J Med* **70:** 517, 520-1, 525-6.

Kim K.J., Sunshine M.G., Lindqvist B.H., and Six E.W. (2001) Capsid size determination in the P2-P4 bacteriophage system: Suppression of *sir* mutations in P2's capsid gene N by supersid mutations in P4's external scaffold gene *sid*. *Virology* **283:** 49-58.

Kimura M., and Chow C.K. (1984) The complete amino acid sequences of ribosomal proteins L17, L27, and S9 from *Bacillus stearothermophilus*. *Eur J Biochem* **139:** 225-234.

Kirby W.M. (1963) Vancomycin therapy of staphylococcal infections. *Antibiot Chemother* **11:** 84-96.

Korolev S., Hsieh J., Gauss G.H., Lohman T.M., and Waksman G. (1997) Major domain swiveling revealed by the crystal structures of complexes of *E. coli rep* helicase bound to single-stranded DNA and ADP. *Cell* **90:** 635-647.

Kostyuchenko V.A., Chipman P.R., Leiman P.G., Arisaka F., Mesyanzhinov V.V., and Rossmann M.G. (2005) The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat Struct Mol Biol* **12:** 810-813.

Kotzin B.L., Leung D.Y., Kappler J., and Marrack P. (1993) Superantigens and their potential role in human disease. *Adv Immunol* **54:** 99-166.

Kwan T., Liu J., DuBow M., Gros P., and Pelletier J. (2005) The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc Natl Acad Sci U S A* **102:** 5174-5179.

Lander G.C., Evilevitch A., Jeembaeva M., Potter C.S., Carragher B., and Johnson J.E. (2008) Bacteriophage lambda stabilization by auxiliary protein gpD: Timing, location, and mechanism of attachment determined by cryo-EM. *Structure* **16:** 1399-1406.

Lander G.C., Khayat R., Li R., Prevelige P.E., Potter C.S., Carragher B., and Johnson J.E. (2009) The P22 tail machine at subnanometer resolution reveals the architecture of an infection conduit. *Structure* **17:** 789-799.

Lata R., Conway J.F., Cheng N., Duda R.L., Hendrix R.W., Wikoff W.R.*, et al.* (2000) Maturation dynamics of a viral capsid: Visualization of transitional intermediate states. *Cell* **100:** 253-263.

Lauber M.A., Running W.E., and Reilly J.P. (2009) *B. subtilis* ribosomal proteins: Structural homology and post-translational modifications. *J Proteome Res* **8:** 4193-4206.

Leiman P.G., Chipman P.R., Kostyuchenko V.A., Mesyanzhinov V.V., and Rossmann M.G. (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118:** 419-429.

Leiman P.G., Arisaka F., van Raaij M.J., Kostyuchenko V.A., Aksyuk A.A., Kanamaru S., and Rossmann M.G. (2010) Morphogenesis of the T4 tail and tail fibers. *Virol J* **7:** 355.

Leung D.Y., Hauk P., Strickland I., Travers J.B., and Norris D.A. (1998) The role of superantigens in human diseases: Therapeutic implications for the treatment of skin diseases. *Br J Dermatol* **139 Suppl 53:** 17-29.

Leung D.Y., Meissner H.C., Fulton D.R., Murray D.L., Kotzin B.L., and Schlievert P.M. (1993) Toxic shock syndrome toxin-secreting *Staphylococcus aureus* in Kawasaki Syndrome. *Lancet* **342:** 1385-1388.

Levin M.E., Hendrix R.W., and Casjens S.R. (1993) A programmed translational frameshift is required for the synthesis of a bacteriophage lambda tail assembly protein. *J Mol Biol* **234:** 124-139.

Lin Y.C., and Peterson M.L. (2010) New insights into the prevention of staphylococcal infections and toxic shock syndrome. *Expert Rev Clin Pharmacol* **3:** 753-767.

Lina G., Bohach G.A., Nair S.P., Hiramatsu K., Jouvin-Marche E., Mariuzza R., and International Nomenclature Committee for Staphylococcal Superantigens. (2004) Standard nomenclature for the superantigens expressed by staphylococcus. *J Infect Dis* **189:** 2334-2336.

Lindqvist B.H., Deho G., and Calendar R. (1993) Mechanisms of genome propagation and helper exploitation by satellite phage P4. *Microbiol Rev* **57:** 683-702.

Lindsay J.A., and Foster S.J. (2001) Zur: A Zn(2+)-responsive regulatory element of *Staphylococcus aureus*. *Microbiology* **147:** 1259-1266.

Lindsay J.A., Ruzin A., Ross H.F., Kurepina N., and Novick R.P. (1998) The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. *Mol Microbiol* **29:** 527-543.

Liu J., Dehbi M., Moeck G., Arhin F., Bauda P., Bergeron D.*, et al.* (2004) Antimicrobial drug discovery through bacteriophage genomics. *Nat Biotechnol* **22:** 185-191.

Lopez R., Garcia J.L., Garcia E., Ronda C., and Garcia P. (1992) Structural analysis and biological significance of the cell wall lytic enzymes of *Streptococcus pneumoniae* and its bacteriophage. *FEMS Microbiol Lett* **79:** 439-447.

Louiria D.B., Kaminski T., and Buchman J. (1961) Vancomycin in severe staphylococcal infections. *Arch Intern Med* **107:** 225-240.

Lowy F.D. (1998) *Staphylococcus aureus* infections. *N Engl J Med* **339:** 520-532.

Lubbers M.W., Waterfield N.R., Beresford T.P., Le Page R.W., and Jarvis A.W. (1995) Sequencing and analysis of the prolate-headed lactococcal bacteriophage c2 genome and identification of the structural genes. *Appl Environ Microbiol* **61:** 4348-4356.

Luczak E., and Rosenfeld A. (1976) Distance on a hexagonal grid. *IEEE Trans Comput* **25:** 532-533.

Luque A., and Reguera D. (2010) The structure of elongated viral capsids. *Biophys J* **98:** 2993-3003.

Maguire B.A., Beniaminov A.D., Ramu H., Mankin A.S., and Zimmermann R.A. (2005) A protein component at the heart of an RNA machine: The importance of protein L27 for the function of the bacterial ribosome. *Mol Cell* **20:** 427-435.

Maiques E., Ubeda C., Tormo M.A., Ferrer M.D., Lasa I., Novick R.P., and Penades J.R. (2007) Role of staphylococcal phage and SaPI integrase in intra- and interspecies SaPI transfer. *J Bacteriol* **189:** 5608-5616.

Marvik O.J., Jacobsen E., Dokland T., and Lindqvist B.H. (1994a) Bacteriophage P2 and P4 morphogenesis: Assembly precedes proteolytic processing of the capsid proteins. *Virology* **205:** 51-65.

Marvik O.J., Sharma P., Dokland T., and Lindqvist B.H. (1994b) Bacteriophage P2 and P4 assembly: Alternative scaffolding proteins regulate capsid size. *Virology* **200:** 702-714.

Marvik O.J., Dokland T., Nokling R.H., Jacobsen E., Larsen T., and Lindqvist B.H. (1995) The capsid size-determining protein Sid forms an external scaffold on phage P4 procapsids. *J Mol Biol* **251:** 59-75.

Medina E., Wieczorek D., Medina E.M., Yang Q., Feiss M., and Catalano C.E. (2010) Assembly and maturation of the bacteriophage lambda procapsid: GpC is the viral protease. *J Mol Biol* **401:** 813-830.

Menestrina G., Serra M.D., and Prevost G. (2001) Mode of action of beta-barrel pore-forming toxins of the staphylococcal alpha-hemolysin family. *Toxicon* **39:** 1661-1672.

Miller E.M., and Nickoloff J.A. (1995) *Escherichia coli* electrotransformation.Microbiology Today, Humana Press pp. 105-113.

Moore S.D., and Prevelige P.E.,Jr. (2002) A P22 scaffold protein mutation increases the robustness of head assembly in the presence of excess portal protein. *J Virol* **76:** 10245-10255.

Morgan G.J., Hatfull G.F., Casjens S., and Hendrix R.W. (2002) Bacteriophage Mu genome sequence: Analysis and comparison with Mu-like prophages in *Haemophilus, neisseria* and *Deinococcus*. *J Mol Biol* **317:** 337-359.

Morillas M., Eberl H., Allain F.H., Glockshuber R., and Kuennemann E. (2008) Novel enzymatic activity derived from the Semliki Forest Virus capsid protein. *J Mol Biol* **376:** 721-735.

Narita S., Kaneko J., Chiba J., Piemont Y., Jarraud S., Etienne J., and Kamio Y. (2001) Phage conversion of panton-valentine leukocidin in *Staphylococcus aureus*: Molecular analysis of a PVL-converting phage, phiSLT. *Gene* **268:** 195-206.

Naughton R., and Mannion S. (2011) Staphylococcal infections, not just skin deep. *Anesth Analg* **112:** 992-3.

Nemecek D., Gilcrease E.B., Kang S., Prevelige P.E.,Jr, Casjens S., and Thomas G.J.,Jr. (2007) Subunit conformations and assembly states of a DNA-translocating motor: The terminase of bacteriophage P22. *J Mol Biol* **374:** 817-836.

Newbould M.J., Malam J., McIllmurray J.M., Morris J.A., Telford D.R., and Barson A.J. (1989) Immunohistological localisation of staphylococcal toxic shock syndrome toxin (TSST-1) antigen in sudden infant death syndrome. *J Clin Pathol* **42:** 935-939.

Nilssen O., Fossdal C.G., Johansen B.V., and Lindqvist B.H. (1996) Bacteriophage P4 capsid-size determination and its relationship to P2 helper interference. *Virology* **219:** 443-452.

Nishiyori A., Sakaguchi M., Kato H., Igarashi H., and Miwa K. (1994) Toxic shock syndrome toxin-secreting *Staphylococcus aureus* in Kawasaki Syndrome. *Lancet* **343:** 299-300.

Novick R.P. (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* **48:** 1429-1449.

Novick R.P. (1991) Genetic systems in staphylococci. *Methods Enzymol* **204:** 587-636.

Novick R.P. (1963) Analysis by transduction of mutations affecting penicillinase formation in *Staphylococcus aureus*. *J Gen Microbiol* **33:** 121-136.

Novick R.P., Christie G.E., and Penades J.R. (2010) The phage-related chromosomal islands of Gram-positive bacteria. *Nat Rev Microbiol* **8:** 541-551.

Nugent T., and Jones D.T. (2009) Transmembrane protein topology prediction using support vector machines. *BMC Bioinformatics* **10:** 159.

O'Brien L., Kerrigan S.W., Kaw G., Hogan M., Penades J., Litt D.*, et al.* (2002) Multiple mechanisms for the activation of human platelet aggregation by *Staphylococcus aureus*: Roles for the clumping factors ClfA and ClfB, the serine-aspartate repeat protein SdrE and protein A. *Mol Microbiol* **44:** 1033-1044.

Ochi K. (1995) Comparative ribosomal protein sequence analyses of a phylogenetically defined genus, *Pseudomonas*, and its relatives. *Int J Syst Bacteriol* **45:** 268-273.

Pappin D.J., Hojrup P., and Bleasby A.J. (1993) Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* **3:** 327-332.

Parent K.N., Khayat R., Tu L.H., Suhanovsky M.M., Cortines J.R., Teschke C.M.*, et al.* (2010) P22 coat protein structures reveal a novel mechanism for capsid maturation: Stability without auxiliary proteins or chemical crosslinks. *Structure* **18:** 390-401.

Pell L.G., Liu A., Edmonds L., Donaldson L.W., Howell P.L., and Davidson A.R. (2009) The X-ray crystal structure of the phage lambda tail terminator protein reveals the biologically relevant hexameric ring structure and demonstrates a conserved mechanism of tail termination among diverse long-tailed phages. *J Mol Biol* **389:** 938-951.

Pereira M.P., D'Elia M.A., Troczynska J., and Brown E.D. (2008) Duplication of teichoic acid biosynthetic genes in *Staphylococcus aureus* leads to functionally redundant poly(ribitol phosphate) polymerases. *J Bacteriol* **190:** 5642-5649.

Peterson P.K., Wilkinson B.J., Kim Y., Schmeling D., and Quie P.G. (1978) Influence of encapsulation on staphylococcal opsonization and phagocytosis by Human Polymorphonuclear Leukocytes. *Infect Immun* **19:** 943-949.

Plisson C., White H.E., Auzat I., Zafarani A., Sao-Jose C., Lhuillier S.*, et al.* (2007) Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. *EMBO J* **26:** 3720-3728.

Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. (2008) Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids. *J Mol Biol* **380:** 465-475.

Prevost G., Cribier B., Couppie P., Petiau P., Supersac G., Finck-Barbancon V.*, et al.* (1995a) Panton-valentine leucocidin and gamma-hemolysin from *Staphylococcus aureus* ATCC 49775 are encoded by distinct genetic loci and have different biological activities. *Infect Immun* **63:** 4121-4129.

Prevost G., Couppie P., Prevost P., Gayet S., Petiau P., Cribier B.*, et al.* (1995b) Epidemiological data on *Staphylococcus aureus* strains producing synergohymenotropic toxins. *J Med Microbiol* **42:** 237-245.

Rao V.B., and Black L.W. (2010) Structure and assembly of bacteriophage T4 head. *Virol J* **7:** 356.

Rawlings N.D., Barrett A.J., and Bateman A. (2010) MEROPS: The peptidase database. *Nucleic Acids Res* **38:** D227-33.

Richardson A.R., Dunman P.M., and Fang F.C. (2006) The nitrosative stress response of *Staphylococcus aureus* is required for resistance to innate immunity. *Mol Microbiol* **61:** 927-939.

Rossmann M.G., Morais M.C., Leiman P.G., and Zhang W. (2005) Combining X-ray crystallography and electron microscopy. *Structure* **13:** 355-362.

Ruzin A., Lindsay J., and Novick R.P. (2001) Molecular genetics of SaPI1--a mobile pathogenicity island in *Staphylococcus aureus. Mol Microbiol* **41:** 365-377.

Salunke D.M., Caspar D.L., and Garcea R.L. (1989) Polymorphism in the assembly of polyomavirus capsid protein VP1. *Biophys J* **56:** 887-900.

Sansom M.S., and Weinstein H. (2000) Hinges, swivels and switches: The role of prolines in signalling via transmembrane alpha-helices. *Trends Pharmacol Sci* **21:** 445-451.

Schmidt H., and Hensel M. (2004a) Pathogenicity islands in bacterial pathogenesis. *Clin Microbiol Rev* **17:** 14-56.

Schmidt K.A., Donegan N.P., Kwan W.A.,Jr, and Cheung A. (2004b) Influences of SigmaB and *agr* on expression of staphylococcal enterotoxin B (*seb*) in *Staphylococcus aureus*. *Can J Microbiol* **50:** 351-360.

Schneierson S.S., Komrad E.L., and Bryer M.S. (1958) Successful treatment with vancomycin of septicemia following mitral commissurotomy caused by antibiotic resistant strain of *Staphylococcus aureus*. *J Mt Sinai Hosp N Y* **25:** 437-443.

Schwab J.H., Brown R.R., Anderle S.K., and Schlievert P.M. (1993) Superantigen can reactivate bacterial cell wall-induced arthritis. *J Immunol* **150:** 4151-4159.

Seidl K., Bischoff M., and Berger-Bachi B. (2008) CcpA mediates the catabolite repression of *tst* in *Staphylococcus aureus*. *Infect Immun* **76:** 5093-5099.

Seidman C.E., Struhl K., Sheen J., and Jessen T. (2001) Introduction of plasmid DNA into cells.Current protocols in molecular biology. Publisher:New York : Greene Pub. Associates ; Wiley-Interscience, c1988- .

Selmer M., Dunham C.M., Murphy F.V.,4th, Weixlbaumer A., Petry S., Kelley A.C.*, et al.* (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313:** 1935-1942.

Sharples G.J., Ingleston S.M., and Lloyd R.G. (1999) Holliday junction processing in bacteria: Insights from the evolutionary conservation of RuvABC, RecG, and RusA. *J Bacteriol* **181:** 5543-5550.

Shore D., Deho G., Tsipis J., and Goldstein R. (1978) Determination of capsid size by satellite bacteriophage P4. *Proc Natl Acad Sci U S A* **75:** 400-404.

Six E.W. (1975) The helper dependence of satellite bacteriophage P4: Which gene functions of bacteriophage P2 are needed by P4? *Virology* **67:** 249-263.

Six E.W., and Lindqvist B.H. (1978) Mutual derepression in the P2-P4 bacteriophage system. *Virology* **87:** 217-230.

Six E.W., and Klug C.A. (1973) Bacteriophage P4: A satellite virus depending on a helper such as prophage P2. *Virology* **51:** 327-344.

Six E.W., Sunshine M.G., Williams J., Haggard-Ljungquist E., and Lindqvist B.H. (1991) Morphopoietic switch mutations of bacteriophage P2. *Virology* **182:** 34-46.

Skinner D., and Keefer C.S. (1941) Significance of bacteremia caused by Staphylococcus aureus: A study of one hundred and twenty two cases and a review of the literature concerned with experimental infection in animals. *Arch Intern Med* **68:** 851-875.

Smith C.B., Noble V., Bensch R., Ahlin P.A., Jacobson J.A., and Latham R.H. (1982) Bacterial flora of the vagina during the menstrual cycle: Findings in users of tampons, napkins, and sea sponges. *Ann Intern Med* **96:** 948-951.

Song L., Hobaugh M.R., Shustak C., Cheley S., Bayley H., and Gouaux J.E. (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* **274:** 1859-1866.

Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids. *J Mol Biol* **405:** 863-876.

Steven A.C., Heymann J.B., Cheng N., Trus B.L., and Conway J.F. (2005) Virus maturation: Dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr Opin Struct Biol* **15:** 227-236.

Subedi A., Ubeda C., Adhikari R.P., Penades J.R., and Novick R.P. (2007) Sequence analysis reveals genetic exchanges and intraspecific spread of SaPI2, a pathogenicity island involved in menstrual toxic shock. *Microbiology* **153:** 3235-3245.

Tallent S. (2007) Identification and characterization of helper phage gene products involved in mobilization of staphylococcal pathogenicity island SaPI1. PhD dissertation Virginia Commonwealth University Richmond VA : AAT 3286529.

Tallent S.M., Langston T.B., Moran R.G., and Christie G.E. (2007) Transducing particles of *Staphylococcus aureus* pathogenicity island SaPI1 are comprised of helper phage-encoded proteins. *J Bacteriol* **189:** 7520-7524.

Tang J., Lander G.C., Olia A., Li R., Casjens S., Prevelige P.,Jr*, et al.* (2011) Peering down the barrel of a bacteriophage portal: The genome packaging and release valve in P22. *Structure* **19**(4): 496-502.

Tang L., and Johnson J.E. (2002) Structural biology of viruses by the combination of electron cryomicroscopy and X-ray crystallography. *Biochemistry* **41:** 11517-11524.

Teschke C.M., and Parent K.N. (2010) 'Let the phage do the work': Using the phage P22 coat protein structures as a framework to understand its folding and assembly mutants. *Virology* **401:** 119-130.

Throup J.P., Zappacosta F., Lunsford R.D., Annan R.S., Carr S.A., Lonsdale J.T.*, et al.* (2001) The *srhSR* gene pair from *Staphylococcus aureus*: Genomic and proteomic approaches to the identification and characterization of gene function. *Biochemistry* **40:** 10392-10401.

Thuman-Commike P.A., Greene B., Malinski J.A., King J., and Chiu W. (1998) Role of the scaffolding protein in P22 procapsid size determination suggested by T = 4 and T = 7 procapsid structures. *Biophys J* **74:** 559-568.

Topf M., Lasker K., Webb B., Wolfson H., Chiu W., and Sali A. (2008) Protein structure fitting and refinement guided by cryo-EM density. *Structure* **16:** 295-307.

Tormo M.A., Ferrer M.D., Maiques E., Ubeda C., Selva L., Lasa I.*, et al.* (2008) *Staphylococcus aureus* pathogenicity island DNA is packaged in particles composed of phage proteins. *J Bacteriol* **190:** 2434-2440.

Tormo-Mas M.A., Mir I., Shrestha A., Tallent S.M., Campoy S., Lasa I.*, et al.* (2010) Moonlighting bacteriophage proteins derepress staphylococcal pathogenicity islands. *Nature* **465:** 779-782.

Ubeda C., Maiques E., Barry P., Matthews A., Tormo M.A., Lasa I.*, et al.* (2008) SaPI mutations affecting replication and transfer and enabling autonomous replication in the absence of helper phage. *Mol Microbiol* **67:** 493-503.

Ubeda C., Maiques E., Tormo M., Campoy S., Lasa I., Barbe J.*, et al.* (2007) SaPI operon I is required for SaPI packaging and is controlled by LexA. *Mol Microbiol* **65:** 41-50.

Ubeda C., Olivarez N.P., Barry P., Wang H., Kong X., Matthews A.*, et al.* (2009) Specificity of staphylococcal phage and SaPI DNA packaging as revealed by integrase and terminase mutations. *Mol Microbiol* **72:** 98-108.

van Driel R., Traub F., and Showe M.K. (1980) Probable localization of the bacteriophage T4 prehead proteinase zymogen in the center of the prehead core. *J Virol* **36:** 220-223.

VandenBergh M.F., Yzerman E.P., van Belkum A., Boelens H.A., Sijmons M., and Verbrugh H.A. (1999) Follow-up of *Staphylococcus aureus* nasal carriage after 8 years: Redefining the persistent carrier state. *J Clin Microbiol* **37:** 3133-3140.

Varga D.T., and White A. (1961) Suppression of nasal, skin, and aerial staphylococci by nasal application of methicillin. *J Clin Invest* **40:** 2209-2214.

Vinga I., Droge A., Stiege A.C., Lurz R., Santos M.A., Daugelavicius R., and Tavares P. (2006) The minor capsid protein gp7 of bacteriophage SPP1 is required for efficient infection of bacillus subtilis. *Mol Microbiol* **61:** 1609-1621.

Voorhees R.M., Weixlbaumer A., Loakes D., Kelley A.C., and Ramakrishnan V. (2009) Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol* **16:** 528-533.

Voyich J.M., Otto M., Mathema B., Braughton K.R., Whitney A.R., Welty D.*, et al.* (2006) Is panton-valentine leukocidin the major virulence determinant in community-associated methicillin-resistant *Staphylococcus aureus* disease? *J Infect Dis* **194:** 1761-1770.

Vybiral D., Takac M., Loessner M., Witte A., von Ahsen U., and Blasi U. (2003) Complete nucleotide sequence and molecular characterization of two lytic *Staphylococcus aureus* phages: 44AHJD and P68. *FEMS Microbiol Lett* **219:** 275-283.

Wang I.N., Smith D.L., and Young R. (2000) Holins: The protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54:** 799-825.

Wang S., Chandramouli P., Butcher S., and Dokland T. (2003) Cleavage leads to expansion of bacteriophage P4 procapsids in vitro. *Virology* **314:** 1-8.

Wang S., Palasingam P., Nokling R.H., Lindqvist B.H., and Dokland T. (2000) In vitro assembly of bacteriophage P4 procapsids from purified capsid and scaffolding proteins. *Virology* **275:** 133-144.

Weigel C., and Seitz H. (2006) Bacteriophage replication modules. *FEMS Microbiol Rev* **30:** 321-381.

White R., Chiba S., Pang T., Dewey J.S., Savva C.G., Holzenburg A.*, et al.* (2011) Holin triggering in real time. *Proc Natl Acad Sci U S A* **108:** 798-803.

Wikoff W.R., Liljas L., Duda R.L., Tsuruta H., Hendrix R.W., and Johnson J.E. (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289:** 2129-2133.

WILSON W.L. (1959) The use of vancomycin in staphylococcal infections. *Antibiotic Med Clin Ther* **6:** 167-172.

Witte W., Kresken M., Braulke C., and Cuny C. (1997) Increasing incidence and widespread dissemination of methicillin-resistant *Staphylococcus aureus* (MRSA) in hospitals in central Europe, with special reference to German hospitals. *Clin Microbiol Infect* **3:** 414-422.

Wower I.K., Wower J., and Zimmermann R.A. (1998) Ribosomal protein L27 participates in both 50 S subunit assembly and the peptidyl transferase reaction. *J Biol Chem* **273:** 19847-19852.

Xu J., Hendrix R.W., and Duda R.L. (2004) Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol Cell* **16:** 11-21.

Yamada K., Yanagihara K., Hara Y., Araki N., Harada Y., Morinaga Y.*, et al.* (2011) Clinical features of bacteremia caused by methicillin-resistant *Staphylococcus aureus* in a tertiary hospital. *Tohoku J Exp Med* **224:** 61-67.

Yang K., and Baines J.D. (2009a) Tryptophan residues in the portal protein of Herpes Simplex Virus 1 critical to the interaction with scaffold proteins and incorporation of the portal into capsids. *J Virol* **83:** 11726-11733.

Yang K., and Baines J.D. (2009b) Proline and tyrosine residues in scaffold proteins of Herpes Simplex Virus 1 critical to the interaction with portal protein and its incorporation into capsids. *J Virol* **83:** 8076-8081.

Yarwood J.M., McCormick J.K., and Schlievert P.M. (2001) Identification of a novel two-component regulatory system that acts in global regulation of virulence factors of *Staphylococcus aureus*. *J Bacteriol* **183:** 1113-1123.

Yates J.R.,3rd, Speicher S., Griffin P.R., and Hunkapiller T. (1993) Peptide mass maps: A highly informative approach to protein identification. *Anal Biochem* **214:** 397-408.

Zhang Y. (2009) I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins* **77 Suppl 9:** 100-113.

Zhu W., Clark N.C., McDougal L.K., Hageman J., McDonald L.C., and Patel J.B. (2008) Vancomycin-resistant *Staphylococcus aureus* isolates associated with Inc18-like *vanA* plasmids in Michigan. *Antimicrob Agents Chemother* **52:** 452-457.

Ziebandt A.K., Becher D., Ohlsen K., Hacker J., Hecker M., and Engelmann S. (2004) The influence of *agr* and SigmaB in growth phase dependent regulation of virulence factors in *Staphylococcus aureus*. *Proteomics* **4:** 3034-3047.

Zou D., Kaneko J., Narita S., and Kamio Y. (2000) Prophage, phiPV83-pro, carrying panton-valentine leukocidin genes, on the *Staphylococcus aureus* P83 chromosome: Comparative analysis of the genome structures of phiPV83-pro, phiPVL, phi11, and other phages. *Biosci Biotechnol Biochem* **64:** 2631-2643.

VITA

Priyadarshan Kamalkant Damle was born on March 23rd, 1967 in Mumbai, Maharashtra, India. He is citizen of India. He graduated from Fertilizer's Higher Secondary School, Bharuch, India in 1984. He received his Bachelor of Science degree in chemistry from South Gujarat University in 1987 and Master of Science from South Gujarat University in 1990 with organic chemistry as major and concentration in medicinal chemistry.

Priyadarshan worked for 12 years in a number of Pharmaceutical organizations in various capacities. While working on rabies vaccine at Chiron Behring, he developed interest in viruses. This interest led him to Virginia Commonwealth University where he earned a Doctorate degree in 2011. He hopes to pursue a career in pharmaceutical industries.

Publications:

Poliakov A., Chang J.R., Spilman M.S., Damle P.K., Christie G.E., Mobley J.A., and Dokland T. (2008) Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids. *J Mol Biol* **380:** 465-475.

Spilman M.S., Dearborn A.D., Chang J.R., Damle P.K., Christie G.E., and Dokland T. (2011) A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80 alpha capsids. *J Mol Biol* **405:** 863-876.

Dearborn A.D., Spilman M.S., Damle P.K., Chang J.R., Monroe E.B., Saad J.S., Christie G.E., and Dokland T. (2011)The *Staphylococcus aureus* pathogenicity island 1 protein gp6 functions as an internal scaffold during capsid size determination. *J Mol Biol* **In Press**